## REMARKS

Claims 1 and 13-20 are pending and under consideration. Claims 1 and 13 to 20 have been amended herein. Claims 11 and 12 have been canceled herein without prejudice or disclaimer. Claims 21 and 22 have been added. It is noteworthy that the claim amendments herein are for clarification, to correct typographical errors, or to comply with sequence rules. Therefore, none of the amendments presented herein are narrowing amendments. Upon entry of the present Amendment, claims 1 and 13-22 will be pending and under consideration. It is noteworthy that the Office Action acknowledges that claims 1 and 13-18 appear allowable over the art. (See Office Action page 4, last paragraph).

The amendments and newly added claims do not add new matter. The amendment to claim 1 corrects typographical errors. The amendment to claim 13 is supported, for example, by Table 3 on page 12. The amendment to claims 14-16 substitute sequence identifiers of sequences that correspond to the recited accession numbers. As evidenced by the attached Genbank listings (Exhibit A), the sequences added into the specification herein are identical to the Genbank sequences that were submitted as of the filing date of the present application. All of the Genbank sequences were submitted on June 15, 2001, which predates the November 16, 2001 filing date of the present application, and all of the Genbank entries are version 1 (i.e., The Genbank version ends in ".1"). The amendment to claim 17, which clarifies that the actinomycete is obtained from sediment, is supported, for example, by paragraph [0020]. The amendments to claim 18 correct typographical errors and clarify that the biomolecule is produced by the marine actinomycete, as supported, for example, by paragraph [0013]. The amendments to claims 19 and 20 correct claim dependency, such that these claims properly depend from a method claim (i.e. claim 18). Newly added claims 21 and 22 are supported, for example, by paragraph [0029]. The amendments to the specification add sequence identifiers and correct obvious typographical errors. In paragraph [0032] the correct Genbank accession number, AY040623, is indicated. This number was readily identifiable by a skilled artisan, for example, because the number included in the application as filed, AY040632, is very similar to the correct number, and a search of Genbank for "*Salinospora*" and "CNH964" identifies 2

Genbank entries, only one of which, AY040623, is a 16S rRNA sequence (See Exhibit G). Applicants respectfully request entry of the amendments and reconsideration of the application in view of the amendments and remarks herein.

**Specification Objections**

The Office Action asserts that names of genera, classes, or families of microorganisms should be capitalized and italicized or underlined, such as "Actinomycetes." Applicants respectfully assert that microorganisms and there higher order classifications are correctly indicated in the present specification. As indicated in paragraph [0003], members of the class Actinobacteria are commonly called actinomycetes. As will be understood, common names are not typically capitalized. Therefore, it is proper that "actinomycetes" is not capitalized.

The Office Action asserts that attempts to incorporate sequences into the application by reference to sequences deposited in Genbank are improper. The specification as amended, includes sequence information of referenced Genbank sequences in a substitute sequence listing, and refers to the sequences in the body of the specification with sequence identifiers. Furthermore, submitted herewith is the requisite Statements indicating that the Substitute Sequence Listing does not introduce new matter. As evidenced by the attached Genbank listings, the sequences added into the specification herein are identical to the Genbank sequences that were submitted as of the filing date of the present application. All of the Genbank sequences were submitted on June 15, 2001, which predates the November 16, 2001 filing date of the present application, and all of the Genbank entries are version 1 (i.e., The Genbank version ends in ".1"). Accordingly, Applicants respectfully request withdrawal of the objections to the specification.

**Claim Rejection Under 35 U.S.C. §112, First Paragraph**

Claims 1 and 13-18 stand rejected under 35 U.S.C. §112, first paragraph, as allegedly claiming subject matter that was not described in the specification. Applicants respectfully

traverse the rejection. The Office Action alleges that claim 1 is not adequately described in the specification. The Office Action alleges that there is no basis or support for 18S rRNA, as recited in claim 1 with reference to position 468. Claim 1 as amended, does not recite 18S rRNA. Rather, amended claim 1 recites that position 468 refers to 16S rRNA. Therefore, Applicants respectfully request withdrawal of the rejection of claims 1 and 13-18 under 35 U.S.C. §112, first paragraph.

**Claim Rejection Under 35 U.S.C. §112, Second Paragraph**

Claims 1 and 13-20 stand rejected under 35 U.S.C. §112, second paragraph, as allegedly claiming subject matter that was not adequately described in the specification. Applicants respectfully traverse the rejection. The Office Action alleges that claim 1 is vague and indefinite in the recitation of certain positions in the 16S rRNA without an indication of the context of these positions. The Office Action alleges that the intended sequences must be in the specification and/or claims in proper form for examination.

Applicants respectively assert that a skilled artisan will recognize the precise location of the recited nucleotide positions within the 16S rRNA without further disclosure of 16S rRNA sequences. 16S rRNA is a highly conserved and well-characterized polynucleotide that is commonly used to identify microbes including those of the genera Micromonosporacae (See e.g., Koch et al. *J. System. Bacteriol.*, 46:765-768 (1996), of record, Exhibit B; and present specification paragraph [0027]). As indicated in the specification, the numbering used is in reference to the well characterized *E. coli* numbering system (paragraph [0026]). The procedure of aligning a sequence to secondary structures obtained from a public database is a standard method known in the art to create a "molecular ruler" by which signature nucleotides are reported. Using these well-known sequence alignment tools, a skilled artisan, for example, will identify the signature nucleotides of SEQ ID NOs:3-8 (See Exhibit F (Signature nucleotide bolded and highlighted)).

A database of 16S rRNA sequences and specific methods for sequence alignment of 16S rRNA sequences are available on the internet at the Ribosomal Database Project internet site

rpd.cme.msu.edu. Exhibit D contains pages from the Ribosomal Database Project internet site that provide details regarding alignment of 16S rRNA (See e.g. Section 2.4 "How Sequence Match Works"). The Ribosomal Database Project was available on the internet before the November 16, 2000, priority date of the present application (See e.g., Exhibit E Maidak et al., *Nucl. Acid. Res.* 24:82-85 (1996)). Furthermore, as illustrated in Koch et al. (1996) and Stackebrandt et al. (*J. System. Bacteriol.*, 47:479-491 (1997) (Exhibit C, previously of record in this case)), it is well known in the art that actinomycetes can be identified by identifying signature nucleotides that occur at identified positions of the 16S rRNA. Therefore, a skilled artisan can determine precisely which nucleotides of a 16S rRNA sequence are recited in the pending claims, and can determine whether a 16S rRNA sequence from an unknown organism has the recited signature nucleotides at the recited positions. Accordingly, Applicants request withdrawal of the rejection of claims 1 and 13-20 under 35 U.S.C. §112, second paragraph.

The Office Action alleges that claims 14-16 are indefinite for recitation of Genbank entry numbers without providing the sequences. Furthermore, the Office Action alleges that the claims do not comply with the sequence rules. The specification and claims as amended, include sequence identifiers which identify specific sequences from the Substitute Sequence Listing. The Substitute Sequence Listing, including paper copy and CRF, and the requisite Statements indicating that the Substitute Sequence Listing does not introduce new matter, have been submitted herewith. As evidenced by the attached Genbank listings (Exhibit A), the sequences added into the specification herein are identical to the Genbank sequences that were referenced in the application as filed, and submitted to Genbank as of the filing date of the present application. All of the Genbank sequences were submitted on June 15, 2001, before the November 16, 2001 filing date of the present application, and all of the Genbank entries are version 1 (i.e., The Genbank version ends in ".1"). Accordingly, Applicants request withdrawal of the rejection of claims 14-16 under 35 U.S.C. §112, second paragraph.

The Office Action alleges that claim 17 is confusing in that the meaning of "sediment-derived" is unclear. As suggested by the Examiner, claim 17 as amended indicates that the isolated actinomycete is obtained from sediment. Accordingly, Applicants respectfully request withdrawal of the rejection of claim 17 under 35 U.S.C. §112, second paragraph.

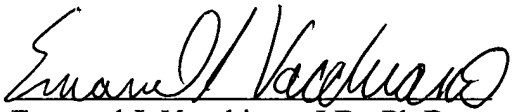
The Office Action alleges that claims 19 and 20 are methods claims that are improperly dependent on a product claim. Furthermore, the Office Action alleges that claim 19 is internally inconsistent in the recitation of "the biomolecules is." As amended, claims 19 and 20 depend from method claim 18. Furthermore, claim 19, recites "the biomolecule is." Accordingly, Applicants respectfully request withdrawal of the rejection of claims 19 and 20 under 35 U.S.C. §112, second paragraph.

The Office Action alleges that claim 18 is vague and indefinite in that it is unclear whether the extracted "biomolecule" is produced by the strain or whether the biomolecule is obtained from the original, uninoculated medium. Claim 18 as amended, clarifies that the biomolecule is produced by the strain. Accordingly, Applicants respectfully request withdrawal of the rejection of claim 18 under 35 U.S.C. §112, second paragraph. In summary, Applicants respectfully request withdrawal of the rejection of claims 1 and 13-20 under 35 U.S.C. §112, second paragraph.

In view of the amendments and the above remarks, it is submitted that the claims are in condition for allowance and a notice to that effect is respectfully requested. The Examiner is invited to contact Applicant's undersigned representative if there are any questions relating to this application. Please charge any additional fees, or make any credits, to Deposit Account No. 50-1355.

Respectfully submitted,

Date: December 15, 2003

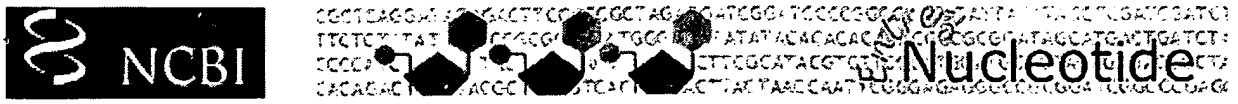


Emanuel J. Vacchiano, J.D., Ph.D.
Registration No. 43,964
Telephone: (858) 638-6754
Facsimile: (858) 677-1465

GRAY CARY WARE & FREIDENRICH LLP
4365 Executive Drive, Suite 1100
San Diego, CA 92121-2133
**USPTO Customer Number: 28213**

**EXHIBIT A**

GENBANK SEQUENCE ENTRIES






1: AY040619. Salinospora sp. C...[gi:22474396] Links

```
LOCUS       AY040619                1481 bp    DNA     linear   BCT 23 -AUG-2002
DEFINITION  Salinospora sp. CNH643 16S ri bosomal RNA gene, partial sequence.
ACCESSION   AY040619
VERSION     AY040619.1  GI:22474396
KEYWORDS    .
SOURCE      Salinospora sp. CNH643
  ORGANISM  Salinospora sp. CNH643
            Bacteria; Firmicutes; Actinob acteria; Actinobacteridae;
            Actinomycetales; Micromonospo rineae; Micromonosporaceae;
            Salinospora.
REFERENCE   1  (bases 1 to 1481)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kau ffman,C.A. and Fenical,W.H.
  TITLE     Novel Marine Actinomycetes Pr esent New Opportunity for Drug
            Discovery
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1481)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kau ffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15 -JUN-2001) Marine Chemistry, Scripps Instituti on of
            Oceanography, UCSD, 8602 La J olla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1481
                     /organism="Salinospo ra sp. CNH643"
                     /mol_type="genomic D NA"
                     /strain="CNH643"
                     /db_xref="taxon:1686 97"
     rRNA            <1..>1481
                     /product="16S riboso mal RNA"
BASE COUNT      315 a    366 c    520 g    280 t
ORIGIN
        1 agagtttgat cctggctcag gacgaacgc t ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggc g aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaaaccg g ggctaatacc ggatatgacc atctgtcgca
      181 tggtgggtgg tggaaagatt ttttggctt g ggatgggcta cgcggcctat cagcttgttg
      241 gtggggtgat ggcctaccaa ggcggcgac g ggtagccggc ctgagagggc gaccggccac
      301 actgggactg agacacggcc cagactcct a cgggaggcag cagtggggaa tcttgcacaa
      361 tgggcggaag cctgatgcag cgacgccgc g tgagggatga cggccttcgg gttgtaaacc
      421 tctttcagca gggacgaagc gtttgtgac g gtacctgcag aagaagcgcc ggccaactac
      481 gtgccagcag ccgcggtaag acgtagggc g caagcgttgt ccggatttat tgggcgtaaa
      541 gagctcgtag gcggcttgtc gcgtcgact g tgaaaacccg tggctcaact gcgggcttgc
      601 agtcgatacg ggcaggctag agttcggta g gggagactgg aattcctggt gtagcggtga
      661 aatgcgcaga tatcaggagg aacaccggt g gcgaaggcgg gtctctgggc cgatactgac
      721 gctgaggagc gaaagcgtgg ggagcgaac a ggattagata ccctggtagt ccacgctgta
      781 aacgttgggg cgtaggtgtg gggggcctc t ccggttctct gtgccgcagc taacgcatta
      841 agcgccccgc ctggggagta cggccgcaa g gctaaaactc aaaggaattg acgggggccc
      901 gcacaagcgg cggagcatgc ggattaatt c gatgcaacgc gaagaacctt acctgggttt
```

```
 961 gacatcgccg gaaatccttc agagatggg g ggtccttcgg ggccggtgac aggtggtgca
1021 tggctgtcgt cagctcgtgt cgtgagatg t tgggttaagt cccgcaacga gcgcaaccct
1081 tgttcgatgt tgccagcgcg ttatggcgg g gactcatcga agactgccgg ggtcaactcg
1141 gaggaaggtg gggatgacgt caagtcatc a tgccccttat gtccagggct tcacgcatgt
1201 tacaatggcc ggtacagtgg gctgcgata c cgtgaggtgg agcgaatccc aaaaagccgg
1261 tctcagttcg gatcggggtc tgcaactcg a ccccgtgaag tcggagtcgc tagtaatcgc
1321 agatcagcaa cgctgcggtg aatacgttc c cgggccttgt acacaccgcc cgtcacgtca
1381 cgaaagtcgg caacacccga agccggtgg c ctaacccttg tgggggagc cgtcgaaggt
1441 ggggctggcg attgggacga agtcgtaac a aggtagccgt a
//
```



Oct 2 2003 18:31:01

NCBI

# Sequence Revision History



**Find** (*Accessions, GI numbers* or *Fasta style SeqIds*) AY040619 G

About Entrez

Entrez

Search for Genes
LocusLink provides curated information for human, fruit fly, mouse, rat, and zebrafish

Help|FAQ

Batch Entrez: Upload a file of GI or accession numbers to retrieve proteine or nucleotide sequences

Check sequence revision history

How to create WWW links to Entrez

LinkOut

Cubby

Related resources

BLAST

Reference sequence project

LocusLink

Clusters of orthologous groups

Protein reviews on the web

Show difference in GenBank/GenPept format Show

| Gi | Version | Update Date |
|---|---|---|
| 22474396 | 1 | Sep 27 2002 9:28 |
| 22474396 | 1 | Aug 23 2002 9:02 |

```
LOCUS       AY040619                1481 bp    DNA     linear   BC
LOCUS       AY040619                1481 bp    DNA     linear   BC
DEFINITION  Salinospora sp. CNH643 16S ribosomal RNA gene, partial
ACCESSION   AY040619
VERSION     AY040619.1  GI:22474396
KEYWORDS    .
SOURCE      Salinospora sp. CNH643
  ORGANISM  Salinospora sp. CNH643
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycet
            Micromonosporineae; Micromonosporaceae; Salinospora.
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Micromonosporineae; Micromonosporacea
            Salinospora.
REFERENCE   1  (bases 1 to 1481)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent populations of a major new ma
            actinomycete taxon in ocean sediments
  JOURNAL   Appl. Environ. Microbiol. 68 (10), 5005-5011 (2002)
  MEDLINE   22235406
   PUBMED   12324350
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.
  TITLE     Novel Marine Actinomycetes Present New Opportunity for
            Discovery
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1481)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Insti
            Oceanography, UCSD, 8602 La Jolla Shores Dr., La Jolla,
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1481
                     /organism="Salinospo ra sp. CNH643"
                     /mol_type="genomic D NA"
```

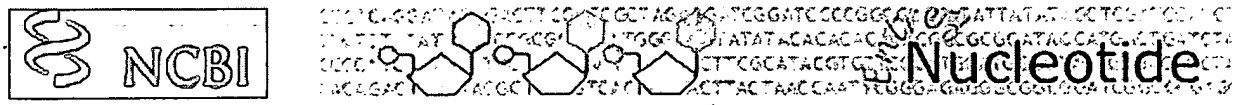
Disclaimer | Write to the Help Desk
NCBI | NLM|NIH




1: AY040620. Salinospora sp. C...[gi:22474393] Links

```
LOCUS       AY040620                1482 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH646 16S ri bosomal RNA gene, partial sequence.
ACCESSION   AY040620
VERSION     AY040620.1  GI:22474393
KEYWORDS    .
SOURCE      Salinospora sp. CNH646
  ORGANISM  Salinospora sp. CNH646
            Bacteria; Actinobacteria; Act inobacteridae; Actinomycetales;
            Micromonosporineae; Micromono sporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1482)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kau ffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent pop ulations of a major new marine
            actinomycete taxon in ocean s ediments
  JOURNAL   Appl. Environ. Microbiol. 68  (10), 5005-5011 (2002)
  MEDLINE   22235406
   PUBMED   12324350
REFERENCE   2  (bases 1 to 1482)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kau ffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Instituti on of
            Oceanography, UCSD, 8602 La J olla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1482
                     /organism="Salinospo ra sp. CNH646"
                     /mol_type="genomic D NA"
                     /strain="CNH646"
                     /db_xref="taxon:1686 98"
     rRNA            <1..>1482
                     /product="16S riboso mal RNA"
BASE COUNT      315 a    369 c    520 g     278 t
ORIGIN
        1 agagtttgat cctggctcag gacgaacgc t ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggc g aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaaaccg g ggctaatacc ggatatgacc atctgtcgca
      181 tggtgggtgg tggaaagatt ttttggctt g ggatgggctc gcggcctatc agcttgttgg
      241 tggggtgatg gcctaccaag gcggcgacg g gtagccggcc tgagagggcg accggccaca
      301 ctgggactga gacacggccc agactccta c gggaggcagc agtggggaat cttgcacaat
      361 gggcggaagc ctgatgcagc gacgccgcg t gagggatgac ggccttcggg ttgtaaacct
      421 ctttcagcag ggacgaagcg tttgtgacg g tacctgcaga agaagcgccg gccaactacg
      481 tgccagcagc cgcggtaaga cgtagggcg c aagcgttgtc cggatttatt gggcgtaaag
      541 agctcgtagg cggcttgtcg cgtcgactg t gaaaacccgt ggctcaactg cgggcttgca
      601 gtcgatacgg gcaggctaga gttcggtag g ggagactgga attcctggtg tagcggtgaa
      661 atgcgcagat atcaggagga acaccggtg g cgaaggcggg tctctgggcc gatactgacg
      721 ctgaggagcg aaagcgtggg gagcgaaca g gattagatac cctggtagtc cacgctgtaa
      781 acgttgggcg ctaggtgtgg ggggcctct c cggttctctg tgccgcagct aacgcattaa
      841 gcgccccgcc tggggagtac ggccgcaag g ctaaaactca aaggaattga cgggggcccg
```

```
     901 cacaagcggc ggagcatgcg gattaattc g atgcaacgcg aagaacctta cctgggtttg
     961 acatcgccgg aaatccttca gagatgggg g gtccttcggg gccggtgaca ggtggtgcat
    1021 ggctgtcgtc agctcgtgtc gtgagatgt t gggttaagtc ccgcaacgag cgcaaccctt
    1081 gttcgatgtt gccagcgcgt tatggcggg g actcatcgaa gactgccggg gtcaactcgg
    1141 aggaaggtgg ggatgacgtc aagtcatca t gccccttatg tccagggctt cacgcatgct
    1201 acaatggccg gtacagtggg ctgcgatacc gtg aggtgga gcgaatccca aaaagccggt
    1261 ctcagttcgg atcggggtct gcaactcga c cccgtgaagt cggagtcgct agtaatcgca
    1321 gatcagcaac gctgcggtga atacgttcc c gggccttgta cacaccgccc gtcacgtcac
    1381 gaaagtcggc aaccaacccg aagccggtg g cctaaccctt gtgggggggag ccgtcgaagg
    1441 gggggctggc gattgggacg aagtcgtaa c aaggtagccg ta
//
```






1: AY040622. Salinospora sp. C...[gi:22474395] Links

```
LOCUS       AY040622                1480 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH898 16S ri bosomal RNA gene, partial sequence.
ACCESSION   AY040622
VERSION     AY040622.1  GI:22474395
KEYWORDS    .
SOURCE      Salinospora sp. CNH898
  ORGANISM  Salinospora sp. CNH898
            Bacteria; Actinobacteria; Act inobacteridae; Actinomycetales;
            Micromonosporineae; Micromono sporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kau ffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent pop ulations of a major new marine
            actinomycete taxon in ocean s ediments
  JOURNAL   Appl. Environ. Microbiol. 68  (10), 5005-5011 (2002)
  MEDLINE   22235406
   PUBMED   12324350
REFERENCE   2  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kau ffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Instituti on of
            Oceanography, UCSD, 8602 La J olla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1480
                     /organism="Salinospo ra sp. CNH898"
                     /mol_type="genomic D NA"
                     /strain="CNH898"
                     /db_xref="taxon:1687 00"
     rRNA            <1..>1480
                     /product="16S riboso mal RNA"
BASE COUNT      316 a    368 c    518 g     278 t
ORIGIN
        1 agagtttgat cctggctcag gacgaacgc t ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggc g aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaaaccg g ggctaatacc ggatatgact ggctgccgca
      181 tggtggttgg tggaaagatt ttttggctt g ggatgggctc gcggcctatc agcttgttgg
      241 tggggtgatg gcctaccaag gcggcgacg g gtagccggcc tgagagggcg accggccaca
      301 ctgggactga gacacggccc agactcctа c gggaggcagc agtggggaat cttgcacaat
      361 gggcggaagc ctgatgcagc gacgccgcg t gagggatgac ggccttcggg ttgtaaacct
      421 ctttcagcag ggacgaagcg tttgtgacg g tacctgcaga agaagcgccg gccaactacg
      481 tgccagcagc cgcggtaaga cgtagggcg c aagcgttgtc cggatttatt gggcgtaaag
      541 agctcgtagg cggcttgtcg cgtcgactg t gaaaacccgt ggctcaactg cgggcttgca
      601 gtcgatacgg gcaggctaga gttcggtag g ggagactgga attcctggtg tagcggtgaa
      661 atgcgcagat atcaggagga acaccggtg g cgaaggcggg tctctgggcc gatactgacg
      721 ctgaggagcg aaagcgtggg gagcgaaca g gattagatac cctggtagtc cacgctgtaa
      781 acgttgggcg ctaggtgtgg ggagcctct c cggttctctg tgccgcagct aacgcattaa
      841 gcgccccgcc tggggagtac ggccgcaag g ctaaaactca aaggaattga cgggggcccg
```

```
      901 cacaagcggc ggagcatgcg gattaattc g atgcaacgcg aagaacctta cctgggtttg
      961 acatcgccgg aaatccttca gagatgggg g gtccttcggg gccggtgaca ggtggtgcat
     1021 ggctgtcgtc agctcgtgtc gtgagatgt t gggttaagtc ccgcaacgag cgcaaccctt
     1081 gttcgatgtt gccagcgcgt tatggcggg g actcatcgaa gactgccggg gtcaactcgg
     1141 aggaaggtgg ggatgacgtc aagtcatca t gccccttatg tccagggctt cacgcatgct
     1201 acaatggccg gtacaatggg ctgcgatacc gtg aggtgga gcgaatccca aaaagccggt
     1261 ctcagttcgg atcggggtct gcaactcga c cccgtgaagt cggagtcgct agtaatcgca
     1321 gatcaacaac gctgcgggga atacgttcc c gggccttgta cacaccgccc gtcacgtcac
     1381 gaaagtcggc aacacccgaa gccggtggc c taacccttgt ggggggagcc gtcgaaggtg
     1441 gggctggcga ttgggacgaa gtcgtaaca a ggtagccgta
//
```




1: AY040617. Salinospora sp. C...[gi:22474391] Links

```
LOCUS       AY040617                1480 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH440 16S ri bosomal RNA gene, partial sequence.
ACCESSION   AY040617
VERSION     AY040617.1  GI:22474391
KEYWORDS    .
SOURCE      Salinospora sp. CNH440
  ORGANISM  Salinospora sp. CNH440
            Bacteria; Actinobacteria; Act inobacteridae; Actinomycetales;
            Micromonosporineae; Micromono sporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kau ffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent pop ulations of a major new marine
            actinomycete taxon in ocean s ediments
  JOURNAL   Appl. Environ. Microbiol. 68  (10), 5005-5011 (2002)
  MEDLINE   22235406
   PUBMED   12324350
REFERENCE   2  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kau ffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Instituti on of
            Oceanography, UCSD, 8602 La J olla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1480
                     /organism="Salinospo ra sp. CNH440"
                     /mol_type="genomic D NA"
                     /strain="CNH440"
                     /db_xref="taxon:1686 95"
     rRNA            <1..>1480
                     /product="16S riboso mal RNA"
BASE COUNT      315 a    368 c    518 g     279 t
ORIGIN
        1 agagtttgat cctggctcag gacgaacgc t ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggc g aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaaaccg g ggctaatacc ggatatgact ggctgccgca
      181 tggtggttgg tggaaagatt ttttggctt g ggatgggctc gcggcctatc agcttgttgg
      241 tggggtgatg gcctaccaag gcggcgacg g gtagccggcc tgagagggcg accggccaca
      301 ctgggactga gacacggccc agactcctа c gggaggcagc agtggggaat cttgcacaat
      361 gggcggaagc ctgatgcagc gacgccgcg t gagggatgac ggccttcggg ttgtaaacct
      421 ctttcagcag ggacgaagcg tttgtgacg g tacctgcaga agaagcgccg gccaactacg
      481 tgccagcagc cgcggtaaga cgtagggcg c aagcgttgtc cggatttatt gggcgtaaag
      541 agctcgtagg cggcttgtcg cgtcgactg t gaaaacccgt ggctcaactg cgggcttgca
      601 gtcgatacgg gcaggctaga gttcggtag g ggagactgga attcctggtg tagcggtgaa
      661 atgcgcagat atcaggagga acaccggtg g cgaaggcggg tctctgggcc gatactgacg
      721 ctgaggagcg aaagcgtggg gagcgaaca g gattagatac cctggtagtc cacgctgtaa
      781 acgttgggcg ctaggtgtgg ggagcctct c cggttctctg tgccgcagct aacgcattaa
      841 gcgccccgcc tggggagtac ggccgcaag g ctaaaactca aaggaattga cgggggcccg
```

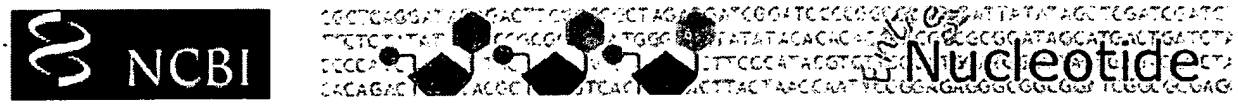
```
     901 cacaagcggc ggagcatgcg gattaattc g atgcaacgcg aagaacctta cctgggtttg
     961 acatcgccgg aaatccttca gagatgggg g gtccttcggg gccggtgaca ggtggtgcat
    1021 ggctgtcgtc agctcgtgtc gtgagatgt t gggttaagtc ccgcaacgag cgcaaccctt
    1081 gttcgatgtt gccagcgcgt tatggcggg g actcatcgaa gactgccggg gtcaactcgg
    1141 aggaaggtgg ggatgacgtc aagtcatca t gccccttatg tccagggctt cacgcatgct
    1201 acaatggccg gtacaatggg ctgcgatacc gtg aggtgga gcgaatccca aaaagccggt
    1261 ctcagttcgg atcggggtct gcaactcga c cccgtgaagt cggagtcgct agtaatcgca
    1321 gatcagcaac gctgcggtga atacgttcc c gggccttgta cacaccgccc gtcacgtcac
    1381 gaaagtcggc aacacccgaa gccggtggc c taacccttgt ggggggagcc gtcgaaggtg
    1441 gggctggcga ttgggacgaa gtcgtaaca a ggtagccgta
//
```




1: AY040618. Salinospora sp. C...[gi:22474392] Links

```
LOCUS       AY040618                1483 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH536 16S ri bosomal RNA gene, partial sequence.
ACCESSION   AY040618
VERSION     AY040618.1  GI:22474392
KEYWORDS    .
SOURCE      Salinospora sp. CNH536
  ORGANISM  Salinospora sp. CNH536
            Bacteria; Actinobacteria; Act inobacteridae; Actinomycetales;
            Micromonosporineae; Micromono sporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1483)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kau ffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent pop ulations of a major new marine
            actinomycete taxon in ocean s ediments
  JOURNAL   Appl. Environ. Microbiol. 68  (10), 5005-5011 (2002)
  MEDLINE   22235406
   PUBMED   12324350
REFERENCE   2  (bases 1 to 1483)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kau ffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Instituti on of
            Oceanography, UCSD, 8602 La J olla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1483
                     /organism="Salinospo ra sp. CNH536"
                     /mol_type="genomic D NA"
                     /strain="CNH536"
                     /db_xref="taxon:1686 96"
     rRNA            <1..>1483
                     /product="16S riboso mal RNA"
BASE COUNT      317 a    369 c    519 g     278 t
ORIGIN
        1 agagtttgat cctggctcag gacgaacgc t ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggc g aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaaaccg g ggctaatacc ggatatgact ggctgccgca
      181 tggtggttgg tggaaagatt ttttggctt g ggatgggctc gcggcctatc agcttgttgg
      241 tggggtgatg gcctaccaag gcggcgacg g gtagccggcc tgagagggcg accggccaca
      301 ctgggactga gacacggccc agactcctа c gggaggcagc agtggggaat cttgcacaat
      361 gggcggaagc ctgatgcagc gacgccgcg t gagggatgac ggccttcggg ttgtaaacct
      421 ctttcagcag ggacggaagc gtttgtgac g gtacctgcag aagaagcgcc ggccaactac
      481 gtgccagcag ccgcggtaag acgtagggc g caagcgttgt ccggaattta ttgggcgtaa
      541 agagctcgta ggcggcttgt cgcgtcgac t gtgaaaaccc gtggctcaac tgcgggcttg
      601 cagtcgatac gggcaggcta gagttcggt a ggggagactg gaattcctgg tgtagcggtg
      661 aaatgcgcag atatcaggag gaacaccgg t ggcgaaggcg ggtctctggg ccgatactga
      721 cgctgaggag cgaaagcgtg gggagcgaa c aggattagat accctggtag tccacgctgt
      781 aaacgttggg cgctaggtgt ggggagcct c tccggttctc tgtgccgcag ctaacgcatt
      841 aagcgccccg cctggggagt acggccgca a ggctaaaact caaaggaatt gacgggggcc
```

```
 901 cgcacaagcg gcggagcatg cggattaat t cgatgcaacg cgaagaacct tacctgggtt
 961 tgacatcgcc ggaaatcctt cagagatgg g gggtccttcg gggccggtga caggtggtgc
1021 atggctgtcg tcagctcgtg tcgtgagat g ttgggttaag tcccgcaacg agcgcaaccc
1081 ttgttcgatg ttgccagcgc gttatggcg g ggactcatcg aagactgccg gggtcaactc
1141 ggaggaaggt ggggatgacg tcaagtcat c atgcccctta tgtccagggc ttcacgcatg
1201 ctacaatggc cggtacaatg ggctgcgata ccg tgaggtg gagcgaatcc caaaaagccg
1261 gtctcagttc ggatcggggt ctgcaactc g accccgtgaa gtcggagtcg ctagtaatcg
1321 cagatcagca acgctgcggt gaatacgtt c ccgggccttg tacacaccgc ccgtcacgtc
1381 acgaaagtcg gcaacacccg aagccggtg g ccctaaccct tgtgggggga gccgtcgaag
1441 gtggggctgg cgaatgggac gaagtcgta a caaggtagcc gta
//
```

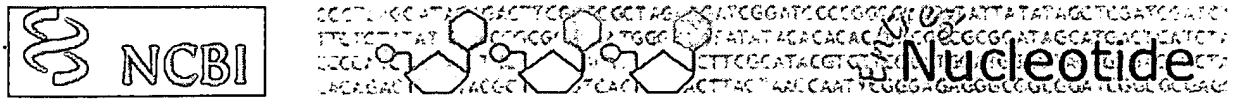





1: AY040621. Salinospora sp. C...[gi:22474394] Links

```
LOCUS       AY040621                1480 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH725 16S ri bosomal RNA gene, partial sequence.
ACCESSION   AY040621
VERSION     AY040621.1  GI:22474394
KEYWORDS    .
SOURCE      Salinospora sp. CNH725
  ORGANISM  Salinospora sp. CNH725
            Bacteria; Actinobacteria; Act inobacteridae; Actinomycetales;
            Micromonosporineae; Micromono sporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kau ffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent pop ulations of a major new marine
            actinomycete taxon in ocean s ediments
  JOURNAL   Appl. Environ. Microbiol. 68  (10), 5005-5011 (2002)
  MEDLINE   22235406
   PUBMED   12324350
REFERENCE   2  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kau ffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Instituti on of
            Oceanography, UCSD, 8602 La J olla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1480
                     /organism="Salinospo ra sp. CNH725"
                     /mol_type="genomic D NA"
                     /strain="CNH725"
                     /db_xref="taxon:1686 99"
     rRNA            <1..>1480
                     /product="16S riboso mal RNA"
BASE COUNT      314 a    369 c    518 g     279 t
ORIGIN
        1 agagtttgat cctggctcag gacgaacgc t ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggc g aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaacccg g ggctaatacc ggatatgacc atctgtcgca
      181 tggtgggtgg tggaaagatt ttttggctt g ggatgggctc gcggcctatc agcttgttgg
      241 tggggtgatg gcctaccaag gcggcgacg g gtagccggcc tgagagggcg accggccaca
      301 ctgggactga gacacggccc agactcctа c gggaggcagc agtggggaat cttgcacaat
      361 gggcggaagc ctgatgcagc gacgccgcg t gagggatgac ggccttcggg ttgtaaacct
      421 ctttcagcag ggacgaagcg tttgtgacg g tacctgcaga aaaagcgccg gccaactacg
      481 tgccagcagc cgcggtaaga cgtagggcg c aagcgttgtc cggatttatt gggcgtaaag
      541 agctcgtagg cggcttgtcg cgtcgactg t gaaaacccgt ggctcaactg cgggcttgca
      601 gtcgatacgg gcaggctaga gttcggtag g ggagactgga attcctggtg tagcggtgaa
      661 atgcgcagat atcaggagga acaccggtg g cgaaggcggg tctctgggcc gatactgacg
      721 ctgaggagcg aaagcgtggg gagcgaaca g gattagatac cctggtagtc cacgctgtaa
      781 acgttgggcg ctaggtgtgg ggggcctct c cggttctctg tgccgcagct aacgcattaa
      841 gcgccccgcc tggggagtac ggccgcaag g ctaaaactca aaggaattga cgggggcccg
```

```
      901 cacaagcggc ggagcatgcg gattaattc g atgcaacgcg aagaacctta cctgggtttg
      961 acatcgccgg aaatccttca gagatgggg g gtccttcggg gccggtgaca ggtggtgcat
     1021 ggctgtcgtc agctcgtgtc gtgagatgt t gggttaagtc ccgcaacgag cgcaaccctt
     1081 gttcgatgtt gccagcgcgt tatggcggg g actcatcgaa gactgccggg gtcaactcgg
     1141 aggaaggtgg ggatgacgtc aagtcatca t gccccttatg tccagggctt cacgcatgct
     1201 acaatggccg gtacagtggg ctgcgatacc gtg aggtgga gcgaatccca aaaagccggt
     1261 ctcagttcgg atcggggtct gcaactcga c cccgtgaagt cggagtcgct agtaatcgca
     1321 gatcagcaac gctgcggtga atacgttcc c gggccttgta cacaccgccc gtcacgtcac
     1381 gaaagtcggc aacacccgaa gccggtggc c taacccttgt ggggggagcc gtcaagtggg
     1441 gggctggcga ttgggacgaa gtcgtaaca a ggtagccgta
//
```






1: AY040623. Salinospora sp. C...[gi:22474397] Links

```
LOCUS       AY040623                1480 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH964 16S ribosomal RNA gene, partial sequence.
ACCESSION   AY040623
VERSION     AY040623.1  GI:22474397
KEYWORDS    .
SOURCE      Salinospora sp. CNH964
  ORGANISM  Salinospora sp. CNH964
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Micromonosporineae; Micromonosporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent populations of a major new marine
            actinomycete taxon in ocean sediments
  JOURNAL   Appl. Environ. Microbiol. 68 (10), 5005-5011 (2002)
  MEDLINE   22235406
   PUBMED   12324350
REFERENCE   2  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Institution of
            Oceanography, UCSD, 8602 La Jolla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1480
                     /organism="Salinospora sp. CNH964"
                     /mol_type="genomic DNA"
                     /strain="CNH964"
                     /db_xref="taxon:168701"
     rRNA            <1..>1480
                     /product="16S ribosomal RNA"
ORIGIN
        1 agagtttgat cctggctcag gacgaacgct ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggcg aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaaaccg g ggctaatacc ggatatgacc atctgtcgca
      181 tggtgggtgg tggaaagatt ttttggcttg ggatgggctc gcggcctatc agcttgttgg
      241 tggggtgatg gcctaccaag gcggcgacgg gtagccggcc tgagagggcg accggccaca
      301 ctgggactga gacacggccc agactcctac gggaggcagc agtggggaat cttgcacaat
      361 gggcggaagc ctgatgcagc gacgccgcg tgagggatgac ggccttcggg ttgtaaacct
      421 ctttcagcag ggacgaagcg tttgtgacgg tacctgcaga agaagcgccg gccaactacg
      481 tgccagcagc cgcggtaaga cgtagggcg caagcgttgtc cggatttatt gggcgtaaag
      541 agctcgtagg cggcttgtcg cgtcgactg tgaaaacctgt ggctcaactg cgggcttgca
      601 gtcgatacgg gcaggctaga gttcggtag ggagactgga attcctggtg tagcggtgaa
      661 atgcgcagat atcaggagga acaccggtg gcgaaggcggg tctctgggcc gatactgacg
      721 ctgaggagcg aaagcgtggg gagcgaaca ggattagatac cctggtagtc cacgctgtaa
      781 acgttgggcg ctaggtgtgg ggggcctct cggttctctg tgccgcagct aacgcattaa
      841 gcgccccgcc tggggagtac ggccgcaag gctaaaactca aaggaattga cgggggcccg
      901 cacaagcggc ggagcatgcg gattaattc gatgcaacgcg aagaacctta cctgggtttg
```

```
     961 acatcgccgg aaatccttca gagatgggg g gtccttcggg gccggtgaca ggtggtgcat
    1021 ggctgtcgtc agctcgtgtc gtgagatgt t gggttaagtc ccgcaacgag cgcaaccctt
    1081 gttcgatgtt gccagcgcgt tatggcggg g actcatcgaa gactgccggg gtcaactcgg
    1141 aggaaggtgg ggatgacgtc aagtcatca t gccccttatg tccagggctt cacgcatgct
    1201 acaatggccg gtacagtggg ctgcgatac c gtgaggtgga gcgaatccca aaaagccggt
    1261 ctcagttcgg atcggggtct gcaactcgac ccc gtgaagt cggagtcgct agtaatcgca
    1321 gataaacaac gctgcggtga atacgttcc c gggccttgta cacaccgccc gtcacgtcac
    1381 gaaagtcggc aacacccgaa gccggtggc c taacccttgt ggggggagcc gtcgaaggtg
    1441 gggctggcga ttgggacgaa gtcgtaaca a ggtagccgta
//
```

**EXHIBIT B**

Copy of Koch et al., *Int. J. Syst. Bacteriol.* 46: 765-768 (1996)

INTERNATIONAL JOURNAL OF SYSTEMATIC BACTERIOLOGY, July 1996, p. 765–768 Vol. 46, No. 3
0020-7713/96/$04.00+0


# 16S Ribosomal DNA Analysis of the Genera *Micromonospora*, *Actinoplanes*, *Catellatospora*, *Catenuloplanes*, *Couchioplanes*, *Dactylosporangium*, and *Pilimelia* and Emendation of the Family *Micromonosporaceae*

CATHRIN KOCH, REINER M. KROPPENSTEDT, FRED A. RAINEY, AND ERKO STACKEBRANDT*

*DSMZ-German Collection of Microorganisms and Cell Cultures GmbH, 38124 Braunschweig, Germany*

**In order to determine the phylogenetic structure of the actinomycete family *Micromonosporaceae*, the 16S ribosomal DNA sequences of 17 type species of the genera *Actinoplanes*, *Dactylosporangium*, and *Pilimelia* were compared with the 16S ribosomal DNA sequences of species of the genera *Catellatospora*, *Catenuloplanes*, and *Couchioplanes* and with those of species of the genus *Micromonospora* and other actinomycete genera for which the sequences have been previously determined. All genera of the family, together with the genera *Catellatospora*, *Catenuloplanes*, and *Couchioplanes*, form a phylogenetically coherent cluster that is well separated from other families of the order *Actinomycetales*. Except for one species of the genus *Catellatospora*, all species cluster according to their genus affiliation. The inclusion of the genera *Catellatospora*, *Catenuloplanes*, and *Couchioplanes* into the family broadens the phenotypic and chemotaxonomic heterogeneity of this taxon. An emendation of the family *Micromonosporaceae* is given.**

An emendation of the original monogeneric family *Micromonosporaceae* Krasilni'kov 1938 (11) has recently lead to an increase in the number of genera with the addition of the genera *Actinoplanes*, *Dactylosporangium*, and *Pilimelia* (6). The genus *Catellatospora* can also be considered a member of this family, as Asano and Kawamoto (1) placed this genus in the actinoplanates group of actinomycetes. The phylogenetic analysis (22) of the recently described genera *Catenuloplanes* (27) and *Couchioplanes* (22) indicated that these genera were so closely related to *Actinoplanes philippinensis* and *Dactylosporangium aurantiacum* that they could also be considered members of the family *Micromonosporaceae*. In contrast to many other actinomycete genera for which the phylogenetic position and taxonomic coherency have been investigated by 16S ribosomal DNA (rDNA) analysis (14, 16, 18, 25, 26), very little information is available on other genera of the family *Micromonosporaceae* (4, 20, 21). Until recently, sequences were available for only 3 species of the about 50 validly described species of this family (22). In a first attempt to elucidate the intrafamily structure of the family *Micromonosporaceae*, the intrageneric phylogenetic structure was determined for the genus *Micromonospora* by analyzing all type strains of the 14 presently available valid species of this genus (9). As a result, the intrageneric coherency of the genus could be confirmed. In this contribution, we extend this survey by presenting 16S rDNA sequence data on members of all genera of the family *Micromonosporaceae*.

## MATERIALS AND METHODS

**Strains investigated and culture conditions.** Strains investigated in this study are listed in Table 1. All strains are deposited in the German Collection of Microorganisms and Cell Cultures GmbH. Except for two strains of the genus *Pilimelia* which were grown on medium 440, all strains were cultivated on medium 65 (*Streptomyces* medium) as indicated in the German Collection of Microorganisms and Cell Cultures catalog of strains (3).

**Analysis of 16S rDNA.** Extraction of genomic DNA and amplification of the 16S rDNA were carried out as described previously (17). PCR products were sequenced directly with the *Taq* DyeDeoxy terminator cycle sequencing kit (Applied Biosystems) according to the manufacturer's protocol. The sequence reactions were electrophoresed with the Applied Biosystems 373A DNA sequencer.

**Phylogenetic analysis.** 16S rDNA sequences were compared with sequences in the existing 16S rDNA database of members of the order *Actinomycetales*. This database consists of sequences deposited in the Ribosomal Database Project (13). Similarity values were transformed into phylogenetic distance values that compensate for multiple substitutions at any given site in the sequence (8). The least-squares distance method of De Soete (2) as well as the programs Neighbor Joining (19) and Maximum Likelihood contained in the PHYLIP package (5) were used in the construction of phylogenetic dendrograms. For the calculation of bootstrap values, 200 trees were analyzed with the programs NJFIND and NJBOOT (kindly provided by T.S. Whittam, Department of Biology, Pennsylvania State University).

**Nucleotide sequence accession numbers.** Sequences have been deposited at EMBL under the accession numbers shown in Table 1.

## RESULTS AND DISCUSSION

The almost complete 16S rDNA sequences (96% of the *Escherichia coli* sequence) of 17 type strains of 16 validly published species from six genera of the family *Micromonosporaceae* were determined and aligned with 14 recently published sequences of type strains of *Micromonospora* species and of *Dactylosporangium thailandense* (9), *Couchioplanes caeruleus* subsp. *caeruleus*, *Catenuloplanes niger*, and *Catenuloplanes atrovinosus* (the sequences are given in reference 22 and the species are described in reference 23). As the sequences of the latter three species contained gaps of significant lengths (the sequences comprise 82% of the *E. coli* sequence), they were analyzed separately from those for which almost complete sequences were available. Homologous sequences from members of the genera *Streptosporangium*, *Streptomyces*, and *Sporichthya* and the family *Pseudonocardiaceae* were used as outgroup sequences. All *Micromonosporaceae* species form a coherent cluster within the actinomycete subphylum of gram-positive bacteria (4), and the distinctness of the family is supported by a bootstrap value of 100% (Fig. 1).

The similarity values for members of the family *Micromonosporaceae* ranged above 94%. This value is on the average 3% higher than those found for members of the family and the outside species (the range is between 1.5 and 5.0%). Similar

* Corresponding author. Mailing address: DSMZ-German Collection of Microorganisms and Cell Cultures GmbH, Mascheroder Weg 1b, 38124 Braunschweig, Germany. Phone: 49 531 2616 352. Fax: 49 531 2616 418. Electronic mail address: erko@gbf-braunschweig.de.

TABLE 1. Strains investigated in this study and 16S rDNA accession numbers

| Taxon | Strain designation | Accession no. |
|---|---|---|
| *Actinoplanes brasiliensis* | DSM 43805$^T$ | X93185 |
| *Actinoplanes cyaneus* | DSM 46137$^T$ | X93186 |
| *Actinoplanes philippinensis* | DSM 43019$^T$ | X93187 |
| *Actinoplanes regularis* | DSM 43151$^T$ | X93188 |
| *Catellatospora citrea* subsp. *citrea* | DSM 44097$^T$ | X93197 |
| *Catellatospora citrea* subsp. *methionotrophica* | DSM 44098$^T$ | X93198 |
| *Catellatospora ferruginea* | DSM 44099$^T$ | X93199 |
| *Catellatospora tsunoense* | DSM 44101$^T$ | X93200 |
| *Catenuloplanes japonicus* | DSM 44102$^T$ | X93201 |
| *Couchioplanes caeruleus* subsp. *azureus* | DSM 44103$^T$ | X93202 |
| *Dactylosporangium aurantiacum* | DSM 43157$^T$ | X93191 |
| *Dactylosporangium fulvum* | DSM 43917$^T$ | X93192 |
| *Dactylosporangium matsuzakiense* | DSM 43810$^T$ | X93193 |
| *Dactylosporangium roseum* | DSM 43916$^T$ | X93194 |
| *"Dactylosporangium salmoneum"* | DSM 43910 | X93195 |
| *Dactylosporangium vinaceum* | DSM 32823$^T$ | X43196 |
| *Pilimelia anulata* | DSM 43039$^T$ | X93189 |
| *Pilimelia terevasa* | DSM 43040$^T$ | X93190 |

intrafamily values have been found for the families *Cellulomonadaceae* (>92% [18]), *Nocardiaceae* (>90% [16], recalculated from recent entries in the public database [13]), and *Mycobacteriaceae* (>94% [14], recalculated), and those for the family *Pseudonocardiaceae* are slightly lower (>90% [26], recalculated]).

Figure 1 is a distance matrix tree generated on the basis of evolutionary distances (matrix not shown). The phylogenetic branching pattern of the 32 species of the family *Micromonosporaceae* was stable with any of the treeing methods used (least-squares method [2], neighbor-joining method [19], and maximum-likelihood method [5]), in that the members of the genera *Actinoplanes*, *Pilimelia*, *Micromonospora*, *Dactylosporangium*, and *Catellatospora* (with the exception of one species of the last genus) formed phylogenetically coherent entities. In agreement with the phylogenetic distinctness of each genus is the distribution of genus-specific chemotaxonomic properties (22) and signature nucleotides that define these five genera (Table 2). However, except for the genera *Pilimelia* (100%), *Catellatospora* (73%), and *Dactylosporangium* (100%), the clustering of species of the other genera is not supported by high bootstrap values.

**The intrageneric relationships. (i) Phylogeny of the sporangiate genera.** Members of the three genera *Pilimelia*, *Actinoplanes*, and *Dactylosporangium* resemble each other in such gross morphological properties as the formation of sporangia with motile spores and generally the absence of an aerial mycelium (10). The shape of the sporangia, however, is an important diagnostic property to differentiate between the genera *Actinoplanes* and *Dactylosporangium*. *Pilimelia* strains can easily be differentiated from those of the other two genera by their slow growth (24).

(a) *Pilimelia*. The two species of the genus *Pilimelia* that were investigated are closely related (98.4% similarity) and form the deepest branching sublineage among all genera of the family *Micromonosporaceae*. Interestingly, members of this genus are the only representatives of the family that possess *N*-acetyl in the glycan moiety of the peptidoglycan (12), while *N*-glycolyl is present in members of all other genera of this family.

(b) *Actinoplanes*. Of the 17 valid species of the sporangiate genus *Actinoplanes* for which partial 16S rDNA data are available (15), four species that represented the two major phylogenetic groups detected from the analysis of a 500-nucleotide-long stretch were selected. These two subgroups, one containing *A. philippinensis* and *Actinoplanes brasiliensis* (97.3% similarity) and the other containing *Actinoplanes cyaneus* and *Actinoplanes regularis* (97.8% similarity), do not exhibit known chemotaxonomic properties that would support the phylogenetic topology. A few 16S rDNA nucleotide signatures, however, that support the formation of two subgroups are present (Table 3). Whether this bifurcation at the generic level is genuine needs to be investigated by sequence analysis of more *Actinoplanes* species.

(c) *Dactylosporangium*. Phylogenetically, *Dactylosporangium* species are more closely related (>98% similarity) than are species of the other genera of the family. The genus has a unique set of 16S rDNA nucleotides which can be used to clearly discriminate members of this genus from those of the other genera.

**(ii) Phylogeny of the nonsporangiate genera.** In contrast to the other members of the family, the genera *Catenuloplanes*, *Couchioplanes*, *Catellatospora*, and *Micromonospora* do not form

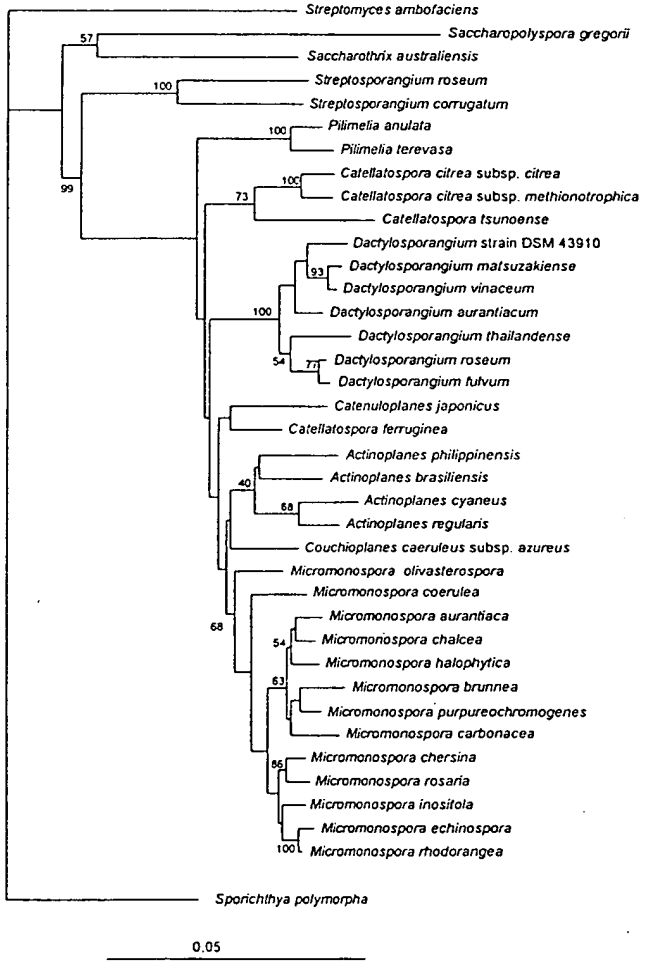



FIG. 1. A distance matrix tree (5) that shows the phylogenetic positions of members of the family *Micromonosporaceae* and some outgroup actinomycetes and that is based on the analysis of almost complete 16S rDNA sequences (dissimilarity matrix not shown). Numbers within the dendrogram indicate the percentages of occurrence in 200 bootstrapped trees. Only values above 40% are shown. The bar represents 5 nucleotide substitutions per 100 nucleotides.

TABLE 2. Distribution of 16S rDNA signature nucleotides which define those genera of the family *Actinoplanaceae* for which significant sequence information is available

| Position(s) in 16S rDNA | Signature nucleotide(s) | | | | |
|---|---|---|---|---|---|
| | *Actino-planes* | *Micromono-spora* | *Dactylo-sporangium* | *Catellato-spora* | *Pilimelia* |
| 1-222 | A-U | G-C | A-U | G-C | G-C |
| 129-132 | U-R[a] | C-R | U-R | U-R | U-R |
| 415 | C | C | A | C | C |
| 441 | A | A | G | A | A |
| 442-492 | G-C | G-C | A · G | G-C | G-C |
| 560 | U | U | U | U | A |
| 600-638 | G-C | G-C | A-U | G-C | G-C |
| 601-637 | A-U | A-U | G-C | U · U | A-U |
| 614-626 | G-C | G-C | G-U | G-C | G-U |
| 653 | A | A | U | A | A |
| 998 | G | G | U | G | U |
| 1992-1041 | G · G | C-G | Var[b] | U-G | G-U |

[a] R, purine.
[b] Var, variable composition.

sporangia. The former two genera produce motile arthrospores on long sporophores, which might be interpreted as a rudimentary aerial mycelium. *Couchioplanes* species possess L-lysine as the diagnostic amino acid in the peptidoglycan (7) rather than *meso*-diaminopimelic acid found in the other species of the family (22, 23, 27). Nevertheless, phylogenetically, they are genuine members of the family *Micromonosporaceae*.

(a) ***Couchioplanes.*** A single species with two subspecies has been described for the genus *Couchioplanes*, i.e., *C. caeruleus* subsp. *caeruleus* (16S rDNA sequenced by Tamura et al. [22]) and *C. caeruleus* subsp. *azureus* (16S rDNA sequenced in this study). On the basis of an analysis of a 1,270-nucleotide-long stretch of the 16S rDNA (omitting the 5′ 120 nucleotides, the 3′ 140 nucleotides, and a 100-nucleotide-long region from positions 920 to 1030 [*E. coli* nomenclature]), the two subspecies share 99% sequence similarity.

(b) ***Catenuloplanes.*** This genus contains six species (23, 27), of which three have fragmentary 16S rDNA sequences only available (22). In order to include the type strain of the type species in the study of the almost complete sequences, the 16S rDNA of *Catenuloplanes japonicus* DSM 44102 (ATCC 31637) was resequenced. The sequence of strain DSM 44102 was 98.8% similar to that of strain ATCC 31637, while lower values were found with *C. niger* IFO 14177T (97.9% similarity) and *C. atrovinosus* RA330 (98.0% similarity). In a phylogenetic analysis in which the fragment length was restricted to the length of the shortest sequence of 1,270 nucleotides (see above), the three species of *Catenuloplanes* clustered together (data not shown) at a position which is identical to that shown for *C. japonicus* in Fig. 1.

(c) ***Catellatospora.*** Members of the genus *Catellatospora* form sporophores from the vegetative mycelium which release nonmotile spores, and an aerial mycelium is absent. Two of the three species investigated, the two subspecies of the type species *Catellatospora citrea* and *Catellatospora tsunoense*, cluster together (96.2% sequence similarity). *Catellatospora ferruginea* branches separately and does not appear to be specifically related to any other species of the family (<97% similarity). The occurrence of two clusters within this genus has already been noticed by Asano and Kawamoto (1). While *C. citrea* and *C. tsunosense* contain menaquinones with 9 isoprene units, *C. ferruginea* and *Catellatospora matsumotoense* (not investigated in this study) contain menaquinones with 10 isoprene units. Whether the latter two species are phylogenetically related and form the nucleus of a novel genus must be discussed when phylogenetic data on *C. matsumotoense* are available.

(d) ***Micromonospora.*** The intrageneric structure of the genus *Micromonospora*, on the basis of the analysis of all type strains and 19 unnamed *Micromonospora* species, has recently been published (9). This genus contains two major subclusters, each of which can be defined by a set of 16S rDNA nucleotide signatures (9).

**Emendation of the family *Micromonosporaceae*.** The family *Micromonosporaceae*, as defined today, contains a collection of four morphologically and chemotaxonomically diverse genera which themselves constitute phenotypically rather distinct entities, i.e., the genera *Micromonospora*, *Actinoplanes*, *Dactylosporangium*, and *Pilimelia* (6). These genera are phylogenetic neighbors, which supports the description of the family *Micromonosporaceae*. As was noted in the description, the family cannot be distinguished from other suprageneric groups of actinomycete families by a set of exclusive phenotypic characters but rather "by DNA homology and rRNA cistron similarities" (6). While the first method does not appear suitable for the delimitation of actinomycete families, the second approach, which has now been replaced by analysis of 16S rDNA sequences, is indeed the only method to allocate new genera to the family. This study shows that the family must be expanded to include the genera *Couchioplanes*, *Catenuloplanes*, and *Catellatospora*. These genera are intermixed with the genera which so far are contained in the family *Micromonosporaceae*. The phenotypic heterogenetity of the family is further increased, especially with respect to the composition of the peptidoglycan, phospholipid type, fatty acid type, and morphology. Tables compiling the differential characteristics have been published by Tamura et al. (22) and Vobis (24).

**Emendation of the family *Micromonosporaceae* Krasil' nikov.** The emended description is based on data compiled by Goodfellow et al. (6), Tamura et al. (22), and Yokota et al. (27) and on data from this study.

Aerobic, gram-positive, non-acid fast organisms that form nonfragmenting, branched, and septate hyphae. Aerial mycelia are rarely developed or scanty. The genus *Micromonospora* produces nonmotile spores. They are born single, sessile, or on short or long sporophores and are spherical, ovoid, or ellipsoidal. The spore wall is thick and may carry blunt spiny ornamentations. In contrast, members of the genera *Actinoplanes*, *Dactylosporangium*, and *Pilimelia* form spores within spore vesicles (sporangia) developed at the tip of short or long sporangiophores. The spore vesicles vary in shape and contain two to five (*Dactylosporangium*) or many (*Actinoplanes* and *Pilimelia*) spores, which may be globose, subglobose, or rod shaped. The spores are motile by means of flagella, which are generally arranged in polar or lateral tufts. In *Couchioplanes* species,

TABLE 3. Distribution of 16S rDNA signature nucleotides which define the two subclusters within the genus *Actinoplanes*

| Position(s) in 16S rDNA | Signature nucleotide(s) | | |
|---|---|---|---|
| | *A. philippinensis* and *A. brasiliensis* | *A. cyaneus* and *A. regularis* | Other genera of the family |
| 138-225 | C-G | U-G | U-G |
| 139-224 | G-U | A-U | A-U |
| 722-733 | G · G | A · A | G · G |
| 1119-1154 | U-A | C/U-G | U-A |
| 1450 | U | G | U |
| 1452 | G | A | G |

spores are formed in chains and are motile. The spore chains and aerial mycelia often aggregate into clusters resembling sporangia, but true sporangia are not observed. In members of the genus *Catenuloplanes*, aerial mycelia are sparse and spores are formed in chains. The spores are rod shaped, straight, or curved with smooth surfaces and are motile by means of peritrichous flagella. *Catellatospora* species lack an aerial mycelium, and short chains of nonmotile spores emerge singly or in tufts from the vegetative hyphae.

The wall peptidoglycan contains *meso*- and/or 3-hydroxydiaminopimelic acid and is of the A1γ type. Except for *Pilimelia* species, which contain acetate as the first amino acid of the peptide chain attached to muramic acid, all other members have glycine; whole-organism hydrolysates are rich in arabinose and xylose, with variable amounts of other sugars. The organisms produce complex mixtures of saturated, iso-, and anteiso-fatty acids and have diphosphatidylglycerol, phosphatidylethanolamine, phosphatidylglycerol, and phosphatidylinositol as major lipids; phosphatidylcholine occurs in addition in *Catenuloplanes* species. Mycolic acids are absent. Menaquinone profiles are heterogeneous. Tetrahydrogenate menaquinones of the MK-9($H_4$) type predominate in members of the genera *Actinoplanes*, *Dactylosporangium*, *Pilimelia*, and *Couchioplanes*, but major amounts of tetra-, hexa-, and/or octahydrogenated menaquinones with 9, 10, and/or 12 isoprene units are found in *Micromonospora* strains. *C. ferruginea* contains MK-10($H_8$) and MK-10($H_6$), while *Catenuloplanes* species contain predominantly MK-9($H_8$) and MK-10($H_8$). The guanine (G)-plus-cytosine (C) content of the DNA is within the range of 71 to 73 mol%. The family can be distinguished from other supragenetic groups of actinomycetes by 16S rDNA similarity.

## REFERENCES


1. **Asano, K., and I. Kawamoto.** 1986. *Catellatospora*, a new genus of the *Actinomycetales*. Int. J. Syst. Bacteriol. 36:512–517.
2. **De Soete, G.** 1983. A least squares algorithm for fitting additive trees to proximity data. Psychometrika 48:621–626.
3. **Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH.** 1993. Catalogue of strains, 5th. ed. Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany.
4. **Embley, T. M., and E. Stackebrandt.** 1994. The molecular phylogeny and systematics of the actinomycetes. Annu. Rev. Microbiol. 48:257–289.
5. **Felsenstein, J.** 1993. Phylip (phylogeny inference package), version 3.5c. University of Washington, Seattle.
6. **Goodfellow, M., L. J. Stanton, K. E. Simpson, and D. E. Minnikin.** 1990. Numerical and chemical classification of *Actinoplanes* and some related actinomycetes. J. Gen. Microbiol. 136:19–36.
7. **Horan, A. C., and B. Brodsky.** 1986. *Actinoplanes caeruleus* sp. nov., a blue-pigmented species of the genus *Actinoplanes*. Int. J. Syst. Bacteriol. 35:187–191.
8. **Jukes, T. H., and C. R. Cantor.** 1969. Evolution of protein molecules, p. 21–132. *In* H. N. Munro (ed.), Mammalian protein metabolism. Academic Press, New York.
9. **Koch, C., R. M. Kroppenstedt, and E. Stackebrandt.** 1996. Intrageneric relationships of the actinomycete genus *Micromonospora*. Int. J. Syst. Bacteriol. 46:383–387.
10. **Kothe, H.-W.** 1987. Die Gattungen *Actinoplanes* und ihre Stellung innerhalb der *Actinomycetales*. Ph.D. thesis. University Marburg, Marburg, Germany.
11. **Krasil'nikov, N. A.** 1938. Ray fungi and related organisms-*Actinomycetales*. Akad. Nauk SSSR, Moscow.
12. **Kroppenstedt, R. M.** 1987. Chemische Untersuchungen an *Actinomycetales* und verwandten Taxa, Korrelation von Chemosystematik und Phylogenie. Habilitation thesis. Technical University Darmstadt, Darmstadt, Germany.
13. **Maidak, B. L., N. Larsen, M. J. McCaughey, R. Overbeck, G. J. Olsen, K. Fogel, J. Blandy, and C. R. Woese.** 1994. The ribosomal database project. Nucleic Acids Res. 22:3483–3487.
14. **Pitulle, C., M. Dorsch, J. Kazda, J. Wolters, and E. Stackebrandt.** 1992. Phylogeny of the rapidly growing members of the genus *Mycobacterium*. Int. J. Syst. Bacteriol. 42:337–343.
15. **Rainey, F. A.** Unpublished data.
16. **Rainey, F. A., J. Burghardt, R. M. Kroppenstedt, S. Klatte, and E. Stackebrandt.** 1995. Phylogenetic analysis of the genera *Rhodococcus* and *Nocardia* and evidence for the evolutionary origin of the genus *Nocardia* from within the radiation of *Rhodococcus* species. Microbiology 141:523–528.
17. **Rainey, F. A., M. Dorsch, H. W. Morgan, and E. Stackebrandt.** 1992. 16S rDNA analysis of *Spirochaeta thermophila*: its phylogenetic position and implications for the systematics of the order *Spirochaetales*. Syst. Appl. Microbiol. 15:197–202.
18. **Rainey, F. A., N. Weiss, and E. Stackebrandt.** 1995. Phylogenetic analysis of the genera *Cellulomonas*, *Promicromonospora*, and *Jonesia* and proposal to exclude the genus *Jonesia* from the family *Cellulomonadaceae*. Int. J. Syst. Bacteriol. 45:649–652.
19. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406–425.
20. **Stackebrandt, E., W. Ludwig, E. Seewaldt, and K.-H. Schleifer.** 1983. Phylogeny of sporeforming members of the order *Actinomycetales*. Int. J. Syst. Bacteriol. 33:173–180.
21. **Stackebrandt, E., and C. R. Woese.** 1981. Towards a phylogeny of actinomycetes and related organisms. Curr. Microbiol. 5:131–136.
22. **Tamura, T., Y. Nakagaito, T. Nishii, T. Hasegawa, E. Stackebrandt, and A. Yokota.** 1994. A new genus of the order *Actinomycetales*, *Couchioplanes* gen. nov., with descriptions of *Couchioplanes caeruleus* (Horan and Brodsky 1986) comb. nov. and *Couchioplanes caeruleus* subsp. *azureus* subsp. nov. Int. J. Syst. Bacteriol. 44:193–203.
23. **Tamura, T., A. Yokota, L. H. Huang, T. Hasegawa, and K. Hatano.** 1995. Five new species of the genus *Catenuloplanes*: *Catenuloplanes niger* sp. nov., *Catenuloplanes indicus* sp. nov., *Catenuloplanes atrovinosus* sp. nov., *Catenuloplanes castaneus* sp. nov., and *Catenuloplanes nepalensis* sp. nov. Int. J. Syst. Bacteriol. 45:858–860.
24. **Vobis, G.** 1991. The genus *Actinoplanes* and related genera, p. 1029–1060. *In* A. Balows, H. G. Trüper, M. Dworkin, W. Harder, and K.-H. Schleifer (ed.), The prokaryotes, 2nd ed. Springer, New York.
25. **Ward-Rainey, N., F. A. Rainey, and E. Stackebrandt.** 1996. The phylogenetic structure of the genus *Streptosporangium*. Syst. Appl. Microbiol. 19:50–55.
26. **Warwick, S., T. Bowen, H. McVeigh, and T. M. Embley.** 1994. A phylogenetic analysis of the family *Pseudonocardiaceae* and the genera *Actinokineospora* and *Saccharothrix* with 16S rRNA sequences and a proposal to combine the genera *Amycolata* and *Pseudonocardia* in an emended genus *Pseudonocardia*. Int. J. Syst. Bacteriol. 44:293–299.
27. **Yokota, A., T. Tamura, T. Hasegawa, and L. H. Huang.** 1993. A new genus of the order *Actinomycetales*: *Catenuloplanes japonicus* gen. nov., sp. nov. Int. J. Syst. Bacteriol. 43:805–812.

**EXHIBIT C**

Coy of Stackebrandt et al. (*J. System. Bacteriol.*, 47:479-491 (1997)

INTERNATIONAL JOURNAL OF SYSTEMATIC BACTERIOLOGY, Apr. 1997, p. 479–491 Vol. 47, No. 2
0020-7713/97/$04.00+0


# Proposal for a New Hierarchic Classification System, *Actinobacteria* classis nov.

ERKO STACKEBRANDT,* FRED A. RAINEY, AND NAOMI L. WARD-RAINEY

*DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, 38124 Braunschweig, Germany*

**A new hierarchic classification structure for the taxa between the taxonomic levels of genus and class is proposed for the actinomycete line of descent as defined by analysis of small subunit (16S) rRNA and genes coding for this molecule (rDNA). While the traditional circumscription of a genus of the actinomycete subphylum is by and large in accord with the 16S rRNA/rDNA-based phylogenetic clustering of these organisms, most of the higher taxa proposed in the past do not take into account the phylogenetic clustering of genera. The rich chemical, morphological and physiological diversity of phylogenetically closely related genera makes the description of families and higher taxa so broad that they become meaningless for the description of the enclosed taxa. Here we present a classification system in which phylogenetically neighboring taxa at the genus level are clustered into families, suborders, orders, subclasses, and a class irrespective of those phenotypic characteristics on which the delineation of taxa has been based in the past. Rather than being based on a listing of a wide array of chemotaxonomic, morphological, and physiological properties, the delineation is based solely on 16S rDNA/rRNA sequence-based phylogenetic clustering and the presence of taxon-specific 16S rDNA/RNA signature nucleotides.**

In their publication "On the nature of global classification," Wheelis et al. (177) based the definition of higher taxa on the molecular level of universally homologous functions. This statement is derived from the high correlation of genealogical trees inferred from several such molecules, e.g., genes coding for 16S rRNA (16S rDNA) (179), 23S rDNA (96), elongation factors involved in the translation process, and the β-subunit of ATPase (97). The authors (177) stress that a basic requirement of a global classification is uniformity in methods and characteristics used in defining and ranking taxa. Nonhomologous characteristics, on the other hand, may be useful in confirming the molecular groupings. Application of this classification strategy led to the description of domains for the three highest taxa recognized today, the *Archaea*, *Bacteria*, and *Eucarya* (180). As a consequence of the description of kingdoms for the major lineages within the domain *Eucarya* (plants, animals, fungi, and protozoa), Woese et al. (180) described the two main lineages within the domain *Archaea* as the kingdom *Crenarchaeota* and the kingdom *Euryarchaeota*.

Within the domain *Bacteria*, more than 15 lineages, which in phylogenetic uniqueness and ancestry are comparable to the archaeal kingdoms, have been identified. The taxonomic rank of kingdom has not yet been proposed for any of these lineages. The taxon class *Proteobacteria* has been proposed for a phylogenetically broad cluster of gram-negative genera, and several orders have been described for some of the phylogenetic lineages that emerged from the comparison of evolutionarily conserved macromolecules, e.g., *Aquificales* (15), *Thermotogales* (67), *Verrucomicrobiales* (173), and *Planctomycetales* (138). These phylogenetically coherent taxa are now used side by side with higher taxa that were described at the beginning of the pre-molecular era, i.e., before or around 1984. While the phylogenetic coherence of the division *Firmacutes* (53), the class *Mollicutes*, and the orders *Chlamydiales*, *Spirochaetales*, and *Myxobacterales* were by and large confirmed following 16S rDNA analyses of their members, the majority of higher taxa represent a collection of phylogenetically diverse families and genera. Examples are the classes *Actinomycetes* (81) and *Photobacteria* (53) and the orders *Clostridiales* and *Bacillales* (119), which need to be redefined in order to make classification consistent with phylogenetic structure.

One of the main lines of descent within the domain *Bacteria* includes a wide range of morphologically diverse organisms, most of which, on the basis of a gram-positive staining reaction, can be considered members of the division *Firmacutes* (53). This lineage comprises organisms with a DNA base composition which generally is above 50 mol% G+C (with a few exceptions) and includes representatives of the class *Actinomycetes* (81), the orders *Actinomycetales* (13) and *Micrococcales* (118), the tribes *Brevibacthereae* and *Micrococceae* (120), and several families of the order *Actinomycetales* as well as additional organisms which were identified as members of this lineage by phylogenetic analyses. This lineage encompasses a wide range of bacteria that irrespective of Gram stain reaction, base composition of DNA, morphology, chemotaxonomic properties, and other characteristics used to delineate bacterial taxa in the past, have a common ancestry (Fig. 1).

The modern era in the classification of organisms that are proposed as members of the class *Actinobacteria* has its origin in three sources: firstly, the establishment of chemotaxonomy that detects differences in the chemical composition of cell constituents such as peptidoglycan, polar lipids and fatty acids, isoprenoid quinones, cytochromes, and the base composition of DNA; secondly, the introduction of DNA-DNA reassociation experiments that measure the gross similarities between single-stranded DNA of strains of closely related species (144); and thirdly, the determination of 16S rRNA and rDNA sequence similarities, which reveals the extent of sequence variation among strains at all levels of relatedness (148). Each of these approaches has contributed to the success of a classification strategy which has been termed polyphasic by Colwell (34).

Although the appropriate methods have been availabl for decades, it took about 30 years to achieve a comprehensive overview of the relatedness among actinomycete bacteria that would allow a proposal for a unified classification system. The

* Corresponding author. Mailing address: DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, 38124 Braunschweig, Germany. Phone: 49 531 2616 352. Fax: 49 531 2616 418. E-mail: erko@gbf-braunschweig.de.

# Class *Actinobacteria*

Subclass *Acidimicrobidae* Order *Acidimicrobiales* Family *Acidimicrobiaceae*
Subclass *Rubrobacteridae* Order *Rubrobacterales* Family *Rubrobacteraceae*
Subclass *Coriobacteridae* Order *Coriobacteriales* Family *Coriobacteriaceae*
Subclass *Sphaerobacteridae* Order *Sphaerobacterales* Family *Sphaerobacteraceae*
Subclass *Actinobacteridae* Order *Actinomycetales*

| Suborder *Actinomycineae* | Suborder *Micrococcineae* | Suborder *Corynebacterineae* | Suborder *Micromonosporineae* | Suborder *Propionibacterineae* |
|---|---|---|---|---|
| Family *Actinomycetaceae* | Families *Micrococcaceae*, *Brevibacteriaceae*, *Cellulomonadaceae*, *Dermabacteraceae*, *Dermatophilaceae*, *Intrasporangiaceae*, *Jonesiaceae*, *Microbacteriaceae*, *Promicromonosporaceae* | Families *Corynebacteriaceae*, *Dietziaceae*, *Gordoniaceae*, *Mycobacteriaceae*, *Nocardiaceae*, *Tsukamurellaceae* | Family *Micromonosporaceae* | Families *Propionibacteriaceae*, *Nocardioidaceae* |

| Suborder *Pseudonocardineae* | Suborder *Streptomycineae* | Suborder *Streptosporangineae* | Suborder *Frankineae* | Suborder *Glycomycineae* |
|---|---|---|---|---|
| Family *Pseudonocardiaceae* | Family *Streptomycetaceae* | Families *Streptosporangiaceae*, *Nocardiopsaceae*, *Thermomonosporaceae* | Families *Frankiaceae*, *Acidothermaceae*, *Geodermatophilaceae*, *Microsphaeraceae*, *Sporichthyaceae* | Family *Glycomycetaceae* |

Order *Bifidobacteriales* Family *Bifidobacteriaceae*

FIG. 2. Proposed hierarchic classification system of the class *Actinobacteria* based on the phylogenetic analyses of the 16S rDNA/rRNA sequence data.

## RESULTS AND DISCUSSION

The basis for the proposal of a novel hierarchic structure (Fig. 2) for the phylogenetically coherent group of actinomycete bacteria and relatives is membership of the same phylogenetic group which was formerly described as a 16S rRNA subdivision or subphylum of the gram-positive bacteria. The common ancestry of the actinomycetes proper and the second subdivision of gram-positive bacteria, defined by clostridia, bacilli, and their relatives, has not yet been convincingly demonstrated by 16S rDNA analyses. However, sequence analyses of glutamine synthetase, glutamate dehydrogenase, and the heat shock protein HSP70 reveal a common ancestry of the gram-positive bacteria (59, 60). The phylogenetic coherence of these organisms supports the description of a kingdom for the *Firmacutes* which would contain two or more classes, one of which embraces the actinomycetes proper. If the results of future studies were to unambiguously demonstrate the common ancestry of members of the two subphyla they could be united under the umbrella of a common higher taxon, the kingdom.

Membership of a new strain to the class *Actinobacteria* is indicated by 16S rDNA sequence similarity values above 80% as determined by comparison of almost-complete 16S rDNA sequences of the new strain and the most deeply branching members of the class, such as *Rubrobacter radiotolerans*, *Acidimicrobium ferrooxidans*, or *Coriobacterium glomerans*, and the presence of signature nucleotides. We are aware that the phylogenetic tree, upon which the conclusions outlined below are based, is a mathematical model of how bacterial evolution occurs. Signature nucleotides are derivatives of the classification process; i.e., signatures are determined for those organisms that are contained within a particular data set. It is also known that a significant increase in species numbers in any of the phylogenetic lineages may lead to a decrease in the number of signatures as the 16S rDNA of more slowly or more rapidly evolving strains may not contain the signature. It is hoped that this proposal will stimulate analyses of other conservative genes from organisms that cluster together by 16S rDNA analyses, so that taxonomic information is provided from additional, independently selected genes and properties and the hierarchic structure proposed here can be tested.

The signatures given below for the higher taxa were chosen for their presence in more than 95% of the members of the respective taxon. The signature pattern for monospecific families must be considered tentative. It should be mentioned that the pattern of signature nucleotides, but not necessarily each individual nucleotide, is indicative of the membership of a taxon to a higher taxon. Bootstrap values (not shown) were determined for the branching points shown in Fig. 3 and were higher than 90% in only a few cases, indicating a lack of statistical significance of the respective branching points. Despite the finding that the majority of branching points are not supported by high bootstrap values, the orders, suborders, and families described previously and below are consistently recoverable from phylogenetic analyses. The lack of high bootstrap

**Class *Actinobacteria* classis nov., Stackebrandt, Rainey, and Ward-Rainey.** *Actinobacteria* (Ac.ti.no.bac.te'ri.a. Gr. n. *actis, actinis*, a ray, beam; Gr. dim. n. *bakterion*, a small rod; *-ia*, proposed ending to denote class; *Actinobacteria*, actinomycete group of bacteria of diverse morphological properties). The class is definable in phylogenetic terms as derived from the analysis of macromolecules of universally homologous functions. Strains of the class *Actinobacteria* can consistently be recovered as members of the same phylogenetic lineage, revealing >80% 16S rDNA/rRNA sequence similarity among each other (Fig. 1), and the presence of the following signature nucleotides in the 16S rDNA/rRNA: an A residue at position 906 and either an A or a C residue at position 955 (except for members of the subclasses *Rubrobacteridae* and *Sphaerobacteridae* [which show U residues at these positions]).

The intraclass relatedness reveals the presence of six phylogenetically distinct lineages which are consistently recovered from phylogenetic analyses (42). These lineages are described as orders (Fig. 3). The 16S rDNA/rRNA signatures defining the higher taxa are based on the available 16S rDNA/rRNA sequences of the type strains of type species of the genera. As certain taxa contain only one or two species, the pattern of signature nucleotides may need to be modified as new species are added to the respective genera.

**Subclass *Acidimicrobidae* subclassis nov., Stackebrandt, Rainey, and Ward-Rainey.** *Acidimicrobidae* (A.ci.di.mi.cro.bi'dae. M.L. n. *Acidimicrobium*, type genus of the subclass; *-idae*, ending to denote a subclass; M.L. fem. pl. n. *Acidimicrobidae*, the *Acidimicrobium* subclass). The subclass contains the type order *Acidimicrobiales*. The 16S rDNA/rRNA signature pattern is as that of the family *Acidimicrobiaceae*.

**Order *Acidimicrobiales* ordo. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Acidomicrobiales* (A.ci.di.mi.cro.bi.a'les. M.L. n. *Acidimicrobium*, type genus of the order; *-ales*, ending to denote an order; M.L. fem. pl. n. *Acidimicrobiales*, the *Acidimicrobium* order). The order contains the type family *Acidimicrobiaceae*. The 16S rDNA/rRNA signature pattern is as that of the family *Acidimicrobiaceae*.

**Family *Acidimicrobiaceae* fam. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Acidimicrobiaceae* (A.ci.di.mi.cro.bi.a'ce.ae. M.L. neut. n. *Acidimicrobium*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Acidimicrobiaceae*, the *Acidimicrobium* family). The following pattern of 16S rDNA/rRNA signature nucleotides and nucleotide pairs defines the family *Acidimicrobiaceae*: 291-309 (U-A), 294-303 (U-A), 408-434 (A-U), 670-736 (C-G), 722-733 (A-C), 955-1225 (C-G), 1118-1155 (C-G), 1311-1326 (A-U), and 1410-1490 (A-U). The family contains the type genus *Acidimicrobium* (24). Phylogenetic analyses have been published previously (17, 23, 149).

**Subclass *Rubrobacteridae* subclassis nov., Rainey, Ward-Rainey, and Stackebrandt.** *Rubrobacteridae* (Ru.bro.bac.te.ri'dae. M.L. masc. n. *Rubrobacter*, type genus of the subclass; *-idae*, ending to denote a subclass; M.L. fem. pl. n. *Rubrobacteridae*, the *Rubrobacter* subclass). The subclass contains the type order *Rubrobacterales*. The 16S rDNA/rRNA signature pattern is as that of the family *Rubrobacteraceae*.

**Order *Rubrobacterales* ordo. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Rubrobacterales* (Ru.bro.bac.te.ra'les. M.L. masc. n. *Rubrobacter*, type genus of the order; *-ales*, ending to denote an order; *Rubrobacterales*, the *Rubrobacter* order). The order contains the type family *Rubrobacteraceae*. The 16S rDNA/rRNA signature pattern is as that of the family *Rubrobacteraceae*.

**Family *Rubrobacteraceae* fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Rubrobacteraceae* (Ru.bro.bac.te.ra'ce.ae. M.L. masc. n. *Rubrobacter*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Rubrobacteraceae*, the *Rubrobacter* family). The following pattern of 16S rDNA/rRNA signature nucleotides and nucleotide pairs defines the family: 127-234 (G-C), 291-309 (U-A), 657-749 (G-C), 681-709 (C-G), 941-1342 (A-U), 955-1225 (U-A), 1051-1207 (C-G), 1115-1185 (C-G), 1311-1326 (A-U), and 1410-1490 (A-U). The family contains the genus *Rubrobacter* (157). A phylogenetic analysis has been published previously (17).

**Subclass *Coriobacteridae* subclassis nov., Stackebrandt, Rainey, and Ward-Rainey.** *Coriobacteridae* (Co.ri.o.bac.te.ri'dae. M.L. neut. n. *Coriobacterium*, type genus of the subclass; *-idae*, ending to denote a subclass; M.L. fem. pl. n. *Coriobacteridae*, the *Coriobacterium* subclass). The subclass contains the type order *Coriobacteriales*. The 16S rDNA/rRNA signature pattern is as that of the family *Coriobacteriaceae*.

**Order *Coriobacteriales* ordo. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Coriobacteriales* (Co.ri.o.bac.te.ri.a'les. M.L. neut. n. *Coriobacterium*, type genus of the order; *-ales*, ending to denote an order; M.L. fem. pl. n. *Coriobacteriales*, the *Coriobacterium* order). The order contains the type family *Coriobacteriaceae*. The 16S rDNA/rRNA signature pattern is as that of the family *Coriobacteriaceae*.

**Family *Coriobacteriaceae* fam. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Coriobacteriaceae* (Co.ri.o.bac.te.ri.a'ce.ae. M.L. neut. n. *Coriobacterium*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Coriobacteriaceae*, the *Coriobacterium* family). The pattern of 16S rDNA/rRNA signature nucleotides of members of the family consists of 113-314 (C-G), 294-303 (G-C), 295-302 (U-A), 670-736 (G-C), 771-808 (U-A), 772-807 (A-U), 823-877 (A-U), 941-1342 (A-U), 950-1231 (U-G), 1120-1153 (G-C), 1148 (C), 1165-1171 (C-G), 1242-1295 (G-C), 1313-1324 (G-C), and 1410-1490 (A-U). The family contains the type genus *Coriobacterium* (61) as well as *Atopobium* (26). Phylogenetic analyses have been published previously (128, 146).

**Subclass *Sphaerobacteridae* subclassis nov., Stackebrandt, Rainey, and Ward-Rainey.** *Sphaerobacteridae* (Sphae.ro.bac.te.ri'dae. M.L. masc. n. *Sphaerobacter*, type genus of the subclass; *-idae*, ending to denote a subclass; M.L. fem. pl. n. *Sphaerobacteridae*, the *Sphaerobacter* subclass). The subclass contains the type order *Sphaerobacterales*. The 16S rDNA/rRNA signature pattern is as that of the family *Sphaerobacteraceae*.

**Order *Sphaerobacterales* ordo. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Sphaerobacterales* (Sphae.ro.bac.te.ra'les. M.L. masc. n. *Sphaerobacter*, type genus of the order; *-ales*, ending to denote an order; M.L. fem. pl. n. *Sphaerobacterales*, the *Sphaerobacter* order). The order contains the type family *Sphaerobacteraceae*. The 16S rDNA/rRNA signature pattern is as that of the family *Sphaerobacteraceae*.

**Family *Sphaerobacteraceae* fam. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Sphaerobacteraceae* (Sphae.ro.bac.te.ra'ce.ae. M.L. masc. n. *Sphaerobacter*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Sphaerobacteraceae*, the *Sphaerobacter* family). The 16S rDNA/rRNA signature pattern for the family consists of 291-309 (U-A), 294-303 (U-A), 408-434 (A-U), 417-426 (C-G), 657-749 (G-C), 670-736 (G-C), 681-709 (C-G), 941-1342 (A-U), 955-1225 (U-A), 1120-1153 (G-C), 1148 (C), and 1351-1371 (C-G). The family contains the type genus *Sphaerobacter* (41). A phylogenetic analysis has been published previously (41).

**Subclass *Actinobacteridae* subclassis nov., Stackebrandt, Rainey, and Ward-Rainey.** *Actinobacteridae* (Ac.ti.no.bac.te.ri'dae. M.L. masc. n. *Actinomyces*, type genus of the subclass; *-idae*, ending to denote a subclass; M.L. fem. pl. n. *Actinobacteridae*, the *Actinomyces* subclass). Members of *Actinobacteri-*

*dae* can consistently be recovered as members of the same phylogenetic lineage (42, 48, 133) (Fig. 3). The subclass contains two orders, *Actinomycetales* and *Bifidobacteriales*. The type order is *Actinomycetales* Buchanan 1917 (13) emend. Members of the subclass contain an insertion of about 100 bases between helices 54 and 55 within domain III of the 23S rDNA (132). This insertion has not been found in members of the subclass *Coriobacteridae* (cf. reference 42). The pattern of 16S rDNA/rRNA signatures consists of nucleotides at positions 291-309 (C-G), 294-303 (C-G), 408-434 (G-C), 670-736 (A-U), 941-1342 (G-C), 1148 (U), and 1410-1490 (G-C).

**Order *Actinomycetales* Buchanan 1917 (13), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Actinomycetales* (Ac.ti.no.my.ce.ta'les. M.L. masc. n. *Actinomyces*, type genus of the order; *-ales*, ending to denote an order; M.L. pl. fem. n. *Actinomycetales*, the *Actinomyces* order). The 16S rDNA signature pattern consists of nucleotides at positions 122-239 (A-G), 449 (A), 450-483 (G-C), 823-877 (G-C), and 1118-1155 (U-A). The order contains the suborders *Actinomycineae*, *Corynebacterineae*, *Frankineae*, *Glycomycineae*, *Micrococcineae*, *Micromonosporineae*, *Propionibacterineae*, *Pseudonocardineae*, *Streptomycineae*, and *Streptosporangineae*. The type suborder is *Actinomycineae*.

**Suborder *Actinomycineae* subordo. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Actinomycineae* (Ac.ti.no.my.ci'ne.ae. M.L. masc. n. *Actinomyces*, type genus of the suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl. n. *Actinomycineae*, the *Actinomyces* suborder). The 16S rDNA signature pattern is as that of the type family *Actinomycetaceae*.

**Family *Actinomycetaceae* Buchanan 1918 (14), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Actinomycetaceae* (Ac.ti.no.my.ce.ta'ce.ae. M.L. masc. n. *Actinomyces*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Actinomycetaceae*, the *Actinomyces* family). The pattern of 16S rDNA signature nucleotides consists of positions 598-640 (U-G), 1059-1198 (U-A), and 1061-1195 (G-U). Genera belonging to the family are the type genus *Actinomyces* (63), *Mobiluncus* (141), and *Arcanobacterium* (32). Phylogenetic analyses have been published previously (51, 87, 99, 142).

**Suborder *Propionibacterineae* subordo. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Propionibacterineae* (Pro.pi.on.i.bac.te.ri'ne.ae. M.L. neut. n. *Propionibacterium*, type genus of the suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl. n. *Propionibacterineae*, the *Propionibacterium* suborder). The pattern of 16S rDNA signatures consists of nucleotides at positions 127-234 (A-U), 603-635 (A-U), 657-749 (G-C), 671-735 (A-U), 986-1219 (U-A), 987-1218 (G-C), 990-1215 (U-G), and 1059-1198 (C-G). The suborder contains the type family *Propionibacteriaceae* and the family *Nocardioidaceae*.

**Family *Propionibacteriaceae* Delwiche 1957 (40), emend. Rainey, Ward-Rainey, and Stackebrandt.** *Propionibacteriaceae* (Pro.pi.o.ni.bac.te.ri.a'ce.ae. M.L. neut. n. *Propionibacterium*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Propionibacteriaceae*, the *Propionibacterium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 66-103 (A-U), 328 (U), 370-391 (C-G), 407-435 (C-G), 602-636 (A-U), 658-748 (A-U), 686 (G), 780 (A), 787 (C), 819 (G), 825-875 (A-U), and 1409-1491 (A-U). Genera included in the family are the type genus *Propionibacterium* (108) as well as *Luteococcus* (163), *Microlunatus* (103), and *Propioniferax* (187). Phylogenetic analyses have been published previously (19, 20, 28, 103, 163, 187).

**Family *Nocardioidaceae* Nesterenko et al. 1985 (104), emend. Rainey, Ward-Rainey, and Stackebrandt.** *Nocardioidaceae* (No.car.di.o.i.da'ce.ae. M.L. masc. n. *Nocardioides*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Nocardioidaceae*, the *Nocardioides* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 66-103 (G-C), 328 (C), 370-391 (G-C), 407-435 (A-U), 602-636 (G-U), 658-748 (U-A), 686 (U), 780 (G), 787 (A), 819 (U), 825-875 (G-C), and 1409-1491 (C-G). Genera included in the family are the type genus *Nocardioides* (116) and *Aeromicrobium* (101). Phylogenetic analyses have been published previously (101, 161).

**Suborder *Micrococcineae* (*Micrococceae* Prevot 1961) (120), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Micrococceae* (Mi.cro.coc.ci'ne.ae. M.L. masc. n. *Micrococcus*, type genus of the suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl. n. *Micrococcineae*, the *Micrococcus* suborder). The pattern of 16S rDNA signatures consists of nucleotides at positions 66-103 (A-U), 70-98 (U-A), 82-87 (G-C), 127-234 (A-U), 449 (A), 598-640 (U-G), 600-638 (U-G), 722-733 (A-A), 952-1229 (C-G), 986-1219 (A-U), 987-1218 (A-U), and 1059-1198 (U-A). The type family is *Micrococcaceae*. Other families in the suborder include *Cellulomonadaceae*, *Promicromonosporaceae*, *Dermatophilaceae*, *Brevibacteriaceae*, *Dermabacteraceae*, *Intrasporangiaceae*, *Jonesiaceae*, and *Microbacteriaceae*.

**Family *Micrococcaceae* Pribham 1929 (121), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Micrococcaceae* (Mi.cro.coc.ca'ce.ae. M.L. masc. n. *Micrococcus*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Micrococcaceae*, the *Micrococcus* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 293-304 (G-U), 610 (G), 598-640 (U-U), 615-625 (G-C), 839-847 (A-U), 859 (U), 1025-1036 (C-G), 1026-1035 (C-G), 1265-1270 (U-G), and 1278 (U). The family contains the type genus *Micrococcus* (25) as well as the genera *Arthrobacter* (35; emended in reference 77), *Kocuria* (150), *Nesterenkonia* (150), *Renibacterium* (137), *Rothia* (52), and *Stomatococcus* (7). Phylogenetic analyses have been published previously (76, 150).

**Family *Cellulomonadaceae* Stackebrandt and Prauser 1991 (147), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Cellulomonadaceae* (Cel.lu.lo.mo.na.da'ce.ae. M.L. fem. n. *Cellulomonas*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Cellulomonadaceae*, the *Cellulomonas* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 30-553 (C-G), 100 (C), 183-194 (A-U), 258-268 (G-C), 610 (A), 615-625 (A-U), 630 (C), 658-748 (G-A), 659-746 (C-G), 694 (G), 747 (C), 832-854 (G-U), 859 (G), 1002-1038 (G-C), 1003-1037 (G-U), 1006-1023 (A-C), and 1256 (C). The family contains the type genus *Cellulomonas* (as emended in references 24 and 153) as well as the genera *Oerskovia* (117; emended in reference 90) and *Rarobacter* (182). Phylogenetic analyses have been published previously (49, 129, 153).

**Family *Promicromonosporaceae* fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Promicromonosporaceae* (Pro.mi.cro.mo.no.spo.ra'ce.ae. M.L. fem. n. *Promicromonospora*, the type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Promicromonosporaceae*, the *Promicromonospora* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 77-92 (G-U), 144-178 (U-G), 183-194 (C-G), 199-218 (U-U), 381 (A), 602-636 (G-U), 630 (C), 658-748 (G), 1002-1038 (Purine-U), 1003-1037 (G-C), 1025-1036 (U-A), and 1267 (U). The family contains the type genus *Promicromonospora* (82). Phylogenetic analyses have been published previously (49, 129).

**Family *Dermatophilaceae* Austwick 1958 (4), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Dermatophilaceae* (Der.ma.to.phi.la'ce.ae. M.L. masc. n. *Dermatophilus*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl.

*Dermatophilaceae*, the *Dermatophilus* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 146-176 (G-U), 153-168 (G-U), 502-543 (G-C), 546 (G), 580-761 (U-A), 602-636 (C-G), 615-625 (G-C), 659-746 (U-A), 825-875 (G-C), 838-848 (Pyr-Pur), and 1251 (G). The family contains the type genus *Dermatophilus* (170) as well as the genera *Kytococcus* (150) and *Dermacoccus* (10). Phylogenetic analyses have been published previously (150, 151).

**Family *Brevibacteriaceae* Breed 1953 (10), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Brevibacteriaceae* (Brev.i.bac.te.ri.a'ce.ae. M.L. neut. n. *Brevibacterium*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Brevibacteriaceae*, the *Brevibacterium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 41-401 (U-A), 69-99 (C-U), 142-221 (U-A), 144-178 (U-G), 407-435 (C-G), 586-755 (U-A), 591-648 (G-U), 612-628 (G-C), 616-624 (C-G), 631 (G), 660-745 (A-U), 670-736 (U-A), 896-903 (U-G), 1011-1018 (U-A), 1012-1017 (G-C), 1244-1293 (U-A), 1254-1283 (A-C), 1256 (A), 1257 (G), 1262 (A), 1263-1272 (C-G), 1310-1327 (U-A), and 1442-1460 (U-G). The family contains the type genus *Brevibacterium* (10; emended in reference 31). Phylogenetic analyses have been published previously (16, 127).

**Family *Dermabacteraceae* fam. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Dermabacteraceae* (Der.ma.bac.te.ra'ce.ae. M.L. masc. n. *Dermabacter*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Dermabacteraceae*, the *Dermabacter* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 153-168 (C-G), 248-276 (U-G), 258-268 (A-U), 280 (U), 407-435 (G-U), 580-761 (U-A), 586-755 (U-A), 589-650 (C-G), 602-636 (C-G), 615-625 (A-U), 838-848 (Pur-Pyr), and 1189 (C). The family contains the type genus *Dermabacter* (69) as well as the genus *Brachybacterium* (27). Phylogenetic analyses have been published previously (16, 139).

**Family *Intrasporangiaceae* fam. nov. Rainey, Ward-Rainey, and Stackebrandt.** *Intrasporangiaceae* (In.tra.spo.ran.gi.a'ce.ae. M.L. neut. n. *Intrasporangium*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Intrasporangiaceae*, the *Intrasporangium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 30-553 (C-G), 69-99 (G-U), 140-223 (G-C), 157-164 (G-C), 258-268 (A-U), 630 (C), 658-748 (G-U), 659-746 (U-A), 660-745 (G-C), 694 (G), 838-848 (C-G), 839-847 (U-A), 859 (C), 1003-1037 (G-C), 1007-1022 (C-G), 1133-1141 (A-U), and 1134-1140 (C-G). The family contains the type genus *Intrasporangium* (71) as well as the genera *Sanguibacter* (47) and *Terrabacter* (29). Phylogenetic analyses have been published previously (29, 47).

**Family *Jonesiaceae* fam. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Jonesiaceae* (Jone.si.a'ce.ae. M.L. fem. n. *Jonesia*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Jonesiaceae*, the *Jonesia* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 153-168 (C-G), 280 (U), 379-384 (G-C), 407-435 (G-U), 445-489 (A-U), 589-650 (U-G), 602-636 (U-G), 615-625 (A-U), 668-738 (U-A), 838-848 (Pur-Pyr), and 1189 (C). The family contains the type genus *Jonesia* (130). Phylogenetic analyses have been published previously (49, 129).

**Family *Microbacteriaceae* Park et al. 1993 (113), emend. Rainey, Ward-Rainey, and Stackebrandt.** *Microbacteriaceae* (Mi.cro.bac.te.ri.a'ce.ae. M.L. masc. n. *Microbacterium*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Microbacteriaceae*, the *Microbacterium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 45-396 (U-A), 144-178 (C-G), 258-268 (A-U), 497 (A), 615-625 (A-U), 694 (G), 771-808 (G-C), 839-847 (G-U), 1256 (C), 1310-1327 (A-U), and 1414-1486 (U-A). The family contains the type genus *Microbacterium* (109) as well as the genera *Agrococcus* (58), *Agromyces* (54), *Aureobacterium* (30), *Clavibacter* (39), *Curtobacterium* (181), and *Rathayibacter* (189). Phylogenetic analyses have been published previously (58, 122, 158, 159).

**Suborder *Corynebacterineae* subordo. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Corynebacterineae* (Co.ry.ne.bac.te.ri'ne.ae. M.L. masc. n. *Corynebacterium*, type genus of the suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl. n. *Corynebacterineae*, the *Corynebacterium* suborder). The pattern of 16S rDNA signatures consists of nucleotides at positions 127-234 (G-C), 131-231 (U-Pur), 502-543 (A-U), 658-748 (A-A), 564 (C), 600-638 (G-C), 601-637 (U-G), 660-745 (U-A), 671-735 (C-G), 819 (G), 952-1229 (U-A), 986-1219 (U-A), 1116-1184 (U-G), and 1414-1486 (U-G). The suborder contains the type family *Corynebacteriaceae* as well as the families *Dietziaceae*, *Gordoniaceae*, *Mycobacteriaceae*, *Nocardiaceae*, and *Tsukamurellaceae*.

**Family *Corynebacteriaceae* Lehmann and Neumann 1907 (95), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Corynebacteriaceae* (Co.ry.ne.bac.te.ri.a'ce.ae. M.L. neut. n. *Corynebacterium*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Corynebacteriaceae*, the *Corynebacterium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 293-304 (G-U), 307 (A), 316-337 (U-G), 468 (U), 508 (U), 586-755 (U-G), 631 (G), 661-744 (G-C), 662-743 (U-G), 771-808 (A-U), 824-876 (C-G), 825-875 (G-C), 837-849 (G-U), 843 (C), and 1059-1198 (U-A). The family contains the type genus *Corynebacterium* (94) as well as the genus *Turicella* (50). Phylogenetic analyses have been published previously (50, 114, 135).

**Family *Mycobacteriaceae* Chester 1897 (21), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Mycobacteriaceae* (My.co.bac.te.ri.a'ce.ae. M.L. neut. n. *Mycobacterium*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Mycobacteriaceae*, the *Mycobacterium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 70-98 (A-U), 293-304 (G-U), 307 (C), 328 (U), 614-626 (A-U), 631 (G), 661-744 (G-C), 824-876 (U-A), 825-875 (A-U), 843 (C), and 1122-1151 (A-U). The family contains the type genus *Mycobacterium* (94). Phylogenetic analyses have been published previously (115, 131).

**Family *Nocardiaceae* Castellani and Chalmers 1919 (18), emend. Rainey, Ward-Rainey, and Stackebrandt.** *Nocardiaceae* (No.car.di.a'ce.ae. M.L. fem. n. *Nocardia* type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Nocardiaceae*, the *Nocardia* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 70-98 (U-A), 139-224 (G-C), 843 (C), 1008-1021 (C-G), 1189 (C), 1244-1293 (C-G), and 1308-1329 (C-G). The family contains the type genus *Nocardia* (169) as well as the genus *Rhodococcus* (190). Phylogenetic analyses have been published previously (22, 124, 134).

**Family *Gordoniaceae* fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Gordoniaceae* (Gor.do.ni.a'ce.ae. M.L. fem. n. *Gordonia*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Gordoniaceae*, the *Gordonia* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 70-98 (A-U), 293-304 (A-U), 307 (U), 661-744 (A-U), 824-876 (U-A), 825-875 (A-U), 843 (U), 1002-1038 (A-U), 1007-1022 (C-G), 1122-1151 (G-C), and 1124-1149 (A-U). The family contains the type genus *Gordonia* (154). Corrigendum: The name *Gordonia*, not *Gordona*, is proposed as the correct etymology. Phylogenetic analyses have been published previously (6, 124, 134).

**Family *Tsukamurellaceae* fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Tsukamurellaceae* (Tsu.ka.mu.rel.la'ce.ae. M.L. fem. n. *Tsukamurella*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Tsukamurellaceae*, the *Tsukamurella* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 70-98 (U-A), 293-304 (G-U), 307 (C), 631 (C), 661-744 (G-C), 824-876 (U-A), 825-875 (A-U), 843 (C), 1007-1022 (G-U), and 1122-1151 (A-U). The family contains the type genus *Tsukamurella* (33). Phylogenetic analyses have been published previously (33, 124, 183).

**Family *Dietziaceae* fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Dietziaceae* (Diet.zi.a'ce.ae. M.L. fem. n. *Dietzia*, type genus of the family; *-aceae*, ending to denote a family. M.L. fem. pl. n. *Dietziaceae*, the *Dietzia* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 70-98 (U-A), 293-304 (G-U), 307 (U), 418-425 (U-A), 508 (U), 614-626 (U-G), 631 (G), 661-744 (A-U), 771-808 (A-U), 824-876 (C-G), 825-875 (G-C), 843 (C), 1049-1198 (U-A), and 1122-1151 (A-U). The family contains the type genus *Dietzia* (125). A phylogenetic analysis has been published previously (124).

**Suborder *Pseudonocardineae* subordo. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Pseudonocardineae* (Pseu.do.no.car'di.ne.ae. M.L. fem. n. *Pseudonocardia*, the type genus of the suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl. n. *Pseudonocardineae*, the *Pseudonocardia* suborder). The pattern of 16S rDNA signatures is as that for the family. The type family is *Pseudonocardiaceae*.

**Family *Pseudonocardiaceae* Warwick et al. 1994 (175), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Pseudonocardiaceae* (Pseu.do.no.car.di.a'ce.ae. M.L. fem. n. *Pseudonocardia*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Pseudonocardiaceae*, the *Pseudonocardia* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 127-234 (G-C), 183-194 (G-U), 502-543 (A-U), 603-635 (C-G), 610 (A), 747 (A), 952-1229 (U-A), 986-1219 (U-A), 987-1218 (G-C), 1001-1039 (Pyr-G), and 1308-1329 (C-G). Comment: Although in phylogenetic terms this family is rather broad, it is currently not possible to subdivide the family due to the lack of an unambiguous pattern of signature nucleotides. The family contains the type genus *Pseudonocardia* (66) as well as the genera *Actinopolyspora* (55), *Actinosynnema* (65), *Amycolatopsis* (92), *Kibdelosporangium* (140), *Kutzneria* (152), *Lentzea* (184), *Saccharomonospora* (106), *Saccharopolyspora* (86), *Saccharothrix* (84), *Streptoalloteichus* (168), and *Thermocrispum* (79). Phylogenetic analyses have been published previously (9, 43, 74, 79, 175, 184).

**Suborder *Streptomycineae* subordo. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Streptomycineae* (Strep.to.my.ci'ne.ae. M.L. masc. n. *Streptomyces*, type genus of the suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl. n. *Streptomycineae*, the *Streptomyces* suborder). The pattern of signature nucleotides of 16S rDNA is as that of the type family *Streptomycetaceae*.

**Family *Streptomycetaceae* Waksman and Henrici 1943 (171), emend. Rainey, Ward-Rainey, and Stackebrandt.** *Streptomycetaceae* (Strep.to.my.ce.ta'ce.ae. M.L. masc. n. *Streptomyces*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Streptomycetaceae*, the *Streptomyces* family). The family is emended to exclude the genus *Sporichthya*. The pattern of 16S rDNA signature nucleotides consists of 71 (G), 80-89 (G-C), 81-88 (C-G), 82-87 (U-G), 127-234 (G-C), 209 (C), 210 (C), 211 (G), 610 (G), 671-735 (U-A), 819 (G), 837-849 (C-G), 950-1231 (U-G), 955-1225 (C-G), 965 (C), 1254-1283 (A-U), and 1409-1491 (C-G). The family contains the type genus *Streptomyces* (171; emended in references 176 and 178). Phylogenetic analyses have been published previo (73, 155, 160, 176, 178).

**Suborder *Streptosporangineae* subordo. nov., Ward-Rai Rainey, and Stackebrandt.** *Streptosporangineae* (Strep.to.s ran.gi'ne.ae. M.L. neut. n. *Streptosporangium*, type genus of suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl *Streptosporangineae*, the *Streptosporangium* suborder). The tern of 16S rDNA signatures consists of nucleotides at p tions 127-234 (A-U), 657-749 (G-Pyr), and 955-1225 (C- The type family is *Streptosporangiaceae*.

**Family *Streptosporangiaceae* Goodfellow et al. 1990 (56) ( idation list no. 34), emend. Ward-Rainey, Rainey, and Stac brandt.** *Streptosporangiaceae* (Strep.to.spo.ran.gi.a'ce.ae. M neut. n. *Streptosporangium*, type genus of the family; *-ace* ending to denote a family; M.L. fem. pl. n. *Streptosporan aceae*, the *Streptosporangium* family). The pattern of 16S rD signatures consists of nucleotides at positions 440-494 (C-( 445-489 (G-C), 501-544 (C-G), 502-543 (G-C), no extra b between positions 453 and 479, 586-755 (U-G), 613-627 (P Pur), 681-709 (U-A), 1116-1184 (U-G), 1137 (U), 1355-13 (A-U), 1436-1465 (GC), and 1422-1478 (G-U). The fam contains the type genus *Streptosporangium* (37) as well as t genera *Herbidospora* (83), *Microbispora* (105), *Microtetraspo* (167), *Planobispora* (164), and *Planomonospora* (166). Phy genetic analyses have been published previously (172, 174).

**Family *Nocardiopsaceae* Rainey et al. 1996 (127), emen Rainey, Ward-Rainey, and Stackebrandt.** *Nocardiopsaceae* (No.car.di.op.sa'ce.ae. M.L. fem. n. *Nocardiopsis*, type genus the family; *-aceae*, ending to denote a family; M.L. fem. pl. *Nocardiopsaceae*, the *Nocardiopsis* family). The pattern of 1 rDNA signatures consists of nucleotides at positions 440-4 (U-G), 442-492 (C-G), 445-489 (C-G), four extra bases b tween positions 453 and 479, 501-544 (G-C), 502-543 (A-U 586-755 (C-G), 603-635 (U-A), 613-627 (C-G), 658-748 (U-A 671-735 (C-G), 681-709 (U-A), 1003-1037 (U-G), 1116-118 (C-G), 1137 (A), 1355-1367 (G-C), 1422-1478 (G-U), an 1435-1466 (A-U). The family contains the type genus *Noca diopsis* (100). A phylogenetic analysis has been published pr viously (127).

**Family *Thermomonosporaceae* fam. nov., Rainey, Ward-Rainey and Stackebrandt.** *Thermomonosporaceae* (Ther.mo.mo.no spo.ra'ce.ae. M.L. fem. n. *Thermomonospora*, type genus of th family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Thermomonosporaceae*, the *Thermomonospora* family). Th pattern of 16S rDNA signatures consists of nucleotides a positions 440-494 (C-G), 442-492 (G-C), four to seven extr bases between position 453 and 479, 501-544 (C-G), 502-54 (G-C), 586-755 (C-G), 613-627 (C-G), 658-748 (C-U), 681-70 (C-G), 1003-1037 (A-G), 1116-1184 (C-G), 1355-1367 (A-U 1422-1478 (G-C), and 1435-1466 (G-C). The family contain the type genus *Thermomonospora* (66) as well as the genera *Actinomadura* (89) and *Spirillospora* (38). Phylogenetic analy ses have been published previously (127, 172).

**Suborder *Micromonosporineae* subordo. nov., Stackebrand Rainey, and Ward-Rainey.** *Micromonosporineae* (Mi.cro.mo no.spo.ri'ne.ae. M.L. fem. n. *Micromonospora*, type genus o the suborder; *-ineae*, ending to denote a suborder; M.L. fem pl. n. *Micromonosporineae*, the *Micromonospora* suborder The pattern of 16S rDNA signature nucleotides is as indicated for the family. The suborder contains the type family *Micromonosporaceae*.

**Family *Micromonosporaceae* Krassilnikov 1938 (80), emen Koch et al. 1996 (78), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Micromonosporaceae* (Mi.cro.mo.no.spo.ra'ce M.L. fem. n. *Micromonospora*, type genus of the family; ending to denote a family; M.L. fem. pl. n. *Micromonospo*

*ceae*, the *Micromonospora* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 66-103 (G-C), 127-234 (A-U), 153-168 (C-G), 502-543 (G-C), 589-650 (C-G), 747 (A), 811 (U), 840-846 (C-G), 952-1229 (C-G), 1116-1184 (C-G), and 1133-1141 (G-C). The family contains the type genus *Micromonospora* (111) as well as the genera *Actinoplanes* (36; emended in reference 145), *Catellatospora* (3), *Couchioplanes* (162), *Catenuloplanes* (185), *Dactylosporangium* (165), and *Pilimelia* (72). Phylogenetic analyses have been published previously (75, 78).

**Suborder *Frankineae* subordo. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Frankineae* (Frank.i'ne.ae. M.L. fem. n. *Frankia*, type genus of the suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl. n. *Frankineae*, the *Frankia* suborder). The pattern of 16S rDNA signatures consists of nucleotides at positions 82-87 (C-G), 127-234 (G-C), 141-222 (G-C), 371-390 (G-C), 502-543 (A-U), and 1003-1037 (G-G). The suborder contains the type family *Frankiaceae* as well as the families *Acidothermaceae*, *Microsphaeraceae*, *Geodermatophilaceae*, and *Sporichthyaceae*. Phylogenetic analyses have been published previously (44, 62, 107, 123, 126).

**Family *Frankiaceae* Becking 1970 (5), emend. Hahn et al. 1989 (62), emend. Normand et al. 1996 (107), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Frankiaceae* (Frank.i.a'ce.ae. M.L. fem. n. *Frankia*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Frankiaceae*, the *Frankia* family). The 16S rDNA signature nucleotide pattern consists of 139-224 (G-C), 148-174 (A-G), 155-166 (U-G), 839-847 (A-G), 987-1218 (G-C), 1059-1198 (C-G), and 1308-1329 (C-G). The family contains the type genus *Frankia* (12). Phylogenetic analyses have been published previously (62, 107).

**Family *Geodermatophilaceae* Normand et al. 1996 (107), emend. Stackebrandt, Rainey, and Ward-Rainey.** *Geodermatophilaceae* (Ge.o.der.ma.to.phi.la'ce.ae. M.L. masc. n. *Geodermatophilus*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Geodermatophilaceae*, the *Geodermatophilus* family. The pattern of 16S rDNA signatures consists of nucleotides at positions 139-224 (C-G), 157-164 (A-U), 158-163 (A-U), 186-191 (C-G), 263 (G), 293-304 (G-U), 986-1219 (U-A), 987-1218 (A-U), 1059-1198 (U-A), and 1308-1329 (U-A). The family contains the type genus *Geodermatophilus* (98) as well as the genus *Blastococcus* (2). Phylogenetic analyses have been published previously (44, 107).

**Family *Microsphaeraceae* fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Microsphaeraceae* (Mi.cro.sphae.ra'ce.ae. M.L. fem. n. *Microsphaera*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Microsphaeraceae*, the *Microsphaera* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 139-224 (C-G), 157-164 (G-C), 186-191 (C-G), 839-847 (U-A), 987-1218 (A-U), 1059-1198 (U-A), and 1308-1329 (U-A). The family contains the type genus *Microsphaera* (188). A phylogenetic analysis has been published previously (188).

**Family *Sporichthyaceae* fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Sporichthyaceae* (Spo.rich.thy.a'ce.ae. M.L. fem. n. *Sporichthya*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Sporichthyaceae*, the *Sporichthya* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 139-224 (U-A), 186-191 (G-C), 600-638 (C-G), 839-847 (U-A), 987-1218 (A-U), 1059-1198 (U-A), and 1308-1329 (U-A). The family contains the type genus *Sporichthya* (91). A phylogenetic analysis has been published previously (126).

**Family *Acidothermaceae* fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Acidothermaceae* (A.ci.do.ther.ma'ce.ae. M.L. masc. n. *Acidothermus*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Acidothermaceae*, the *Acidothermus* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 139-224 (C-G), 186-191 (G-C), 839-847 (A-U), 987-1218 (G-C), 1059-1198 (C-G), and 1308-1329 (C-G). The family contains the type genus *Acidothermus* (102). A phylogenetic analysis has been published previously (123).

**Suborder *Glycomycineae* subordo. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Glycomycineae* (Gly.co.my.ci.ne'ae. M.L. masc. n. *Glycomyces*, type genus of the suborder; *-ineae*, ending to denote a suborder; M.L. fem. pl. n. *Glycomycineae*, the *Glycomyces* suborder). The pattern of signature nucleotides of 16S rDNA is as that of the type family *Glycomycetaceae*.

**Family *Glycomycetaceae* fam. nov., Rainey, Ward-Rainey, and Stackebrandt.** *Glycomycetaceae* (Gly.co.my.ce.ta'ce.ae. M.L. masc. n. *Glycomyces*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Glycomycetaceae*, the *Glycomyces* family). The 16S rDNA pattern of 16S rDNA signature nucleotides contains 70-98 (A-U), 127-234 (G-Pyr), 140-223 (A-U), 229 (G), 366 (U), 415 (C), 449 (C), 534 (G), 681-709 (A-U), 825-875 (G-C), 999-1041 (C-G), 1059-1198 (C-G), 1064-1192 (G-G), 1117-1183 (A-U), and 1309-1328 (C-G). The family contains the type genus *Glycomyces* (85). The phylogenetic position is shown in Fig. 3.

**Order *Bifidobacteriales* ordo. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Bifidobacteriales* (Bi.fi.do.bac.te.ri.a'les. M.L. neut. n. *Bifidobacterium*, type genus of the order; *-ales*, ending to denote an order; M.L. fem. pl. n. *Bifidobacteriales*, the *Bifidobacterium* order). The type family of the order is *Bifidobacteriaceae*. The 16S rDNA nucleotide signature is as that of the family.

**Family *Bifidobacteriaceae* fam. nov., Stackebrandt, Rainey, and Ward-Rainey.** *Bifidobacteriaceae* (Bi.fi.do.bac.te.ri.a'ce.ae. M.L. neut. n. *Bifidobacterium*, type genus of the family; *-aceae*, ending to denote a family; M.L. fem. pl. n. *Bifidobacteriaceae*, the *Bifidobacterium* family). The pattern of 16S rDNA signatures consists of nucleotides at positions 122-239 (G-U), 128-233 (C-G), 450-483 (C-G), 602-636 (G-C), 681-709 (C-G), 688-699 (A-U), 823-877 (A-U), 1118-1155 (C-G), and 1311-1326 (A-U). The family contains the type genus *Bifidobacterium* (110) as well as *Gardnerella* (57). A phylogenetic structure of the family has been published previously (88, 99).

## ACKNOWLEDGMENT

We are indebted to Hans G. Trüper for his advice in the nomenclatural aspect of this work.

## REFERENCES

1. Agre, N. S., and L. N. Guzeva. 1975. New actinomycetes genus *Excellospora* gen. nov. Mikrobiologiya 44:518–523.
2. Ahrens, R., and G. Moll. 1970. Ein neues knospendes Bakterium aus der Ostsee. Arch. Mikrobiol. 70:243–265.
3. Asano, K., and I. Kawamoto. 1986. *Catellatospora*, a new genus of *Actinomycetales*. Int. J. Syst. Bacteriol. 36:512–517.
4. Austwick, P. K. C. 1958. Cutaneous streptothrichosis mycotic dermatitis and strawberry foot rot and the genus *Dermatophilus*. Van Saceghem Vet. Rev. Annot. 4:33–48.
5. Becking, J. H. 1970. *Frankiaceae* fam. nov. *Actinomycetales* with one new combination and six new species of the genus *Frankia* Brunchorst 1887, 174. Int. J. Syst. Bacteriol. 20:201–220.
6. Bendinger, B., F. A. Rainey, R. M. Kroppenstedt, M. Mourmann, and S. Klatte. 1995. *Gordona hydrophobica* sp. nov., isolated from biofilters for waste gas treatment. Int. J. Syst. Bacteriol. 45:544–548.
7. Bergan, T., and M. Kocur. 1982. *Stomatococcus mucilaginosus* gen. nov. sp. nov. ep. rev., a member of the family *Micrococceaeae*. Int. J. Syst. Bacteriol. 32:374–377.
8. Bergey, D. H., F. C. Harrison, R. S. Breed, B. W. Hammer, and F. M. Huntoon. 1923. Bergey's manual of determinative bacteriology, 1st ed. The Williams and Wilkins Co., Baltimore, Md.

**EXHIBIT D**

Portion of the Ribosomal Database Project available on the internet at rpd.cme.msu.edu

MICHIGAN STATE UNIVERSITY

CENTER FOR MICROBIAL ECOLOGY

# Ribosomal Database Project II



## news

10/06/2003, 03:33 pm
September update of the preview release out.
[more]

05/26/2003, 05:12 pm
RDP's poster from ASM 2003 is now available.
[more]

04/03/2003, 11:32 am
New Chimera detection program, Bellerophon, released by Phil Hugenholtz and Thomas Huber.
[more]

11/08/2002, 03:44 pm
Mirror site temporarily down.
[more]

08/08/2002, 03:30 pm
Protocol for improved T-RFLP by capillary electrophoresis available.
[more]

## Welcome to the Ribosomal Database Project

The Ribosomal Database Project (RDP) provides ribosome related data services to the scientific community, including online data analysis, rRNA derived phylogenetic trees, and aligned and annotated rRNA sequences.

***New! Preview Release -- Need Your Feedback!***

Try our new alignment, new Hierarchy Browser, new Sequence Match, and send us feedback on their usability and performance:

Click here to enter the preview site.
(with over 78,000 bacterial sequences)

*Have you tried the RDP Tutorial?*

## online help

*RDP Tutorial!*

## sponsors

National Science Foundation

Office of Biological and Environmental Research

National Institutes of Health

MICHIGAN STATE UNIVERSITY

## mirror site

**Japan**

National Institute of Genetics



*Questions? Mail them to RDP-II Web Support.* [Web Server Statistics]

### 3.1 The Sequence Aligner function:

All comparative analyses of sequence depends upon the accurate alignment of characters. In the case of proteins and nucleic acids, these characters are amino acids and nucleotide bases, respectively. The *Sequence Aligner* function of the RDP aligns submitted sequence(s) from users to the closest matches found in the aligned database.

### 3.2 Accessing the *Sequence Aligner*

As with the previous sections, the first step is to load your sequences. Click the Sequence 2 button below to automatically load the sequence into the text block of the *Sequence Aligner* work page.

Sequence 2 | Sequence 3 | Sequence 4

Scroll down to the Sequence Aligner Options section (Figure 3.1) and request that 10 sequences from the database be included in the results. Finally, click on the submit button and wait for your results.

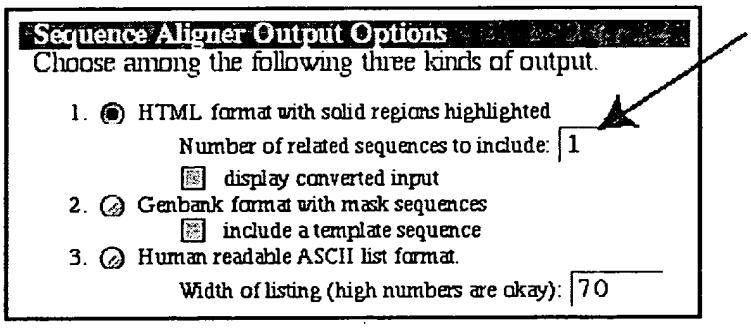



Figure 3.1

### 3.3 Results from Sequence Aligner

The results from your alignment request for Sequence 2 are presented in a browser field that scrolls left and right to accommodate a complete 1500 base sequence (Figure 3.2). The sequence identifiers are links to the genbank files containing the aligned sequence and literature reference (as with the *Sequence Match* function). Scrolling the alignment will reveal regions of the sequence where the Aligner has introduced gaps (--) to bring the sequences into a maximally aligned state. Note that the submitted sequence and the selected aligned sequences from the RDP are highlighted in green. A quick scan of the alignment will convince the reader that Sequence 2 is an exact match to H.inflrrnC, the 16S rRNA from operon C of *Haemophilus influenzae*. This genome of this organism was the first to be sequenced, hence we know the sequence of all six of its rRNA operons (A-F in the alignment). Note there are other *Haemophilus influenza* sequences in the database and that one must look down to sequence 11 before detecting a different species brought into the alignment (*Haemophilus aegyptius*). Sequences 3 and 4 are different in this regard, the related aligned sequences being predominantly from different species.

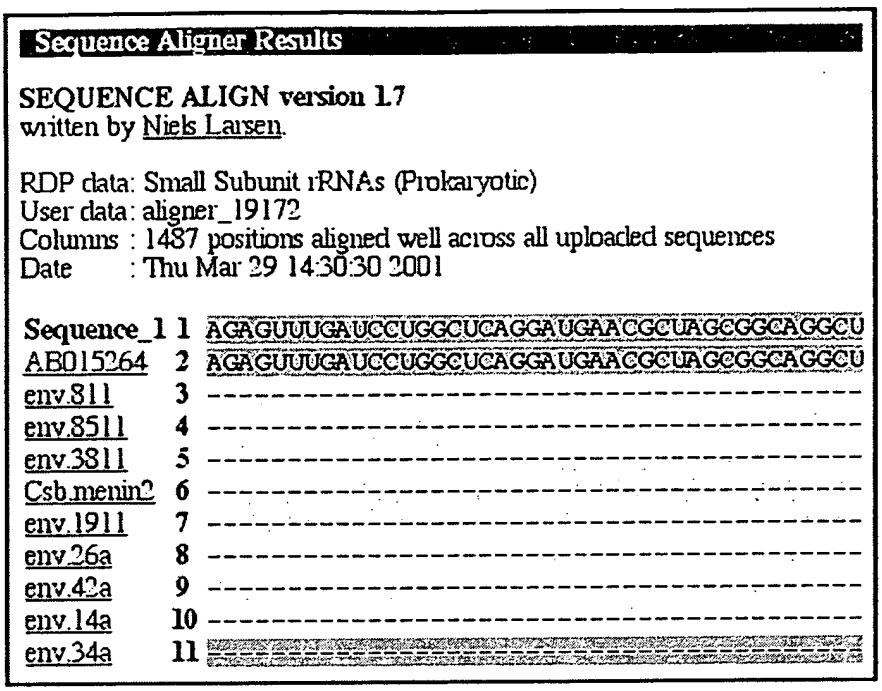



Figure 3.2

### 3.4 How the Sequence Aligner works

The first step in constructing the alignment is to identify the most closely related sequences in the database. This is accomplished using the same approach employed in the *Sequence Match* function except that an eight-base oligomer is used rather than a seven-base oligomer. Hence the submitted sequence(s) is parsed into as many eight base oligomers as possible and each oligomer is used to search the oligomer library of the database. Sequences from the database with the highest score of matching oligomers are retrieved and aligned with the submitted sequence(s). All gaps are removed from the sequences and a pair-wise alignment is made between the submitted sequence and the most closely related sequence from the database, referred to as the template sequence. Gaps are introduced into either of the two sequences to maximize the number of aligned bases. After these two sequences are aligned, the remaining selected sequences are brought into alignment.

### 3.5 *Sequence Aligner* Analysis and Output Options

There are several options associated with this analysis function. Aligner analysis functions (Figure 3.3) include selecting between the Prokaryotic, Eukaryotic or Mitochondrial alignments (these are maintained separately at the RDP), eliminating or preserving common gaps in the alignment (this is probably not an issue for most users), converting T s to U s and vice versa (recall the RDP stores RNA sequences), and finally a choice of case for the display. Note that if you are unsure about which alignment to use, prokaryotic, eukaryotic, or mitochondiral, run your sequence through the *Sequence Match* function and base your alignment selection on the closest match.

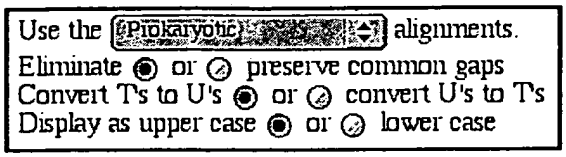



Figure 3.3

There are options for the output display as well (Figure 4.4). These include the format of the output (HTML, Genbank or ASCII), the number of related sequences to include, and the number of characters in the width of the output.

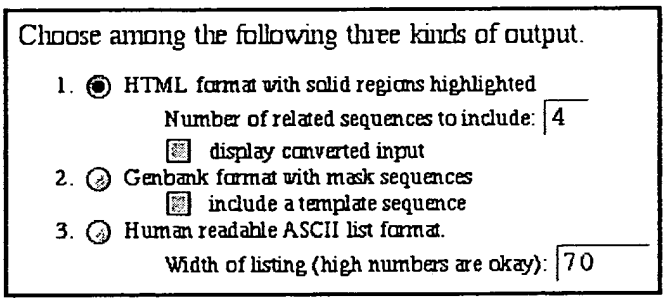



Figure 3.4

### 3.6 The Importance of a Good Alignment

All molecular phylogenetic analyses are based on aligned sequences, be they protein or nucleic acid. The essential tenant of these analyses is that sequences evolve or change over time. Tracking these changes enables us to establish the nature of the relationships between molecules sharing a **common line of descent**. The phrase in bold type is of great importance to us as it is the definition of **homology**, *two or more genes sharing common ancestry*. Hence we are comparing only genes with common ancestry, and similarly, within the gene or protein sequence, we must compare characters of common descent. It is through a multiple sequence alignment that the homology or common ancestry of characters is established. In rRNA alignments, there are regions where the alignment is unamiguous, the sequence is sufficiently conserved such that all characters can be brought into register (alignment) with confidence. There may also be regions with such great variability that no alignment can be made with confidence. The problem with alignments and the concept of homology has resulted in a rich literature. Selected references are included below and a more complete reference list can be downloaded.

### 3.7 The Alignment at the RDP

The rRNA alignment at the RDP is based on our current knowledge of theprimary, secondary, and tertiary structure of the ribosomal RNA molecule and of the ribosome. Users are cautioned that the conservation of the sequence varies from one structural domain to another and that the boundaries of conserved and variable regions of the sequence are not marked. When using the sequences for phylogenetic inference, especially with distance-based methods, a judiciously selected subset of sequence positions should be used. The validity of the inference depends on sequence homology at the positions selected. With Release 8 of the RDP, there are over 16,000 sequences in the prokaryotic alignment, representing the phylogenetic breadth of this group quite well.

## 2.1 What is Sequence Match?

*Sequence Match* is often the first analysis function used by visitors. This function provides a rapid way to screen the entire RDP for the sequences that most closely match a sequence(s) provided by the user. As mentioned in the Chimera Check unit, we provide you with several sequences in order to demonstrate the analyses functions of the RDP. Other sections of the tutorial may use these same sequences. The sequences use in this unit were selected to demonstrate three different possible results. **Sequence_2** has a perfect match within the RDP database, **Sequence_3** has several perfect matches and **Sequence_4** has no perfect matches.

## 2.2 Accessing the Sequence Match Function

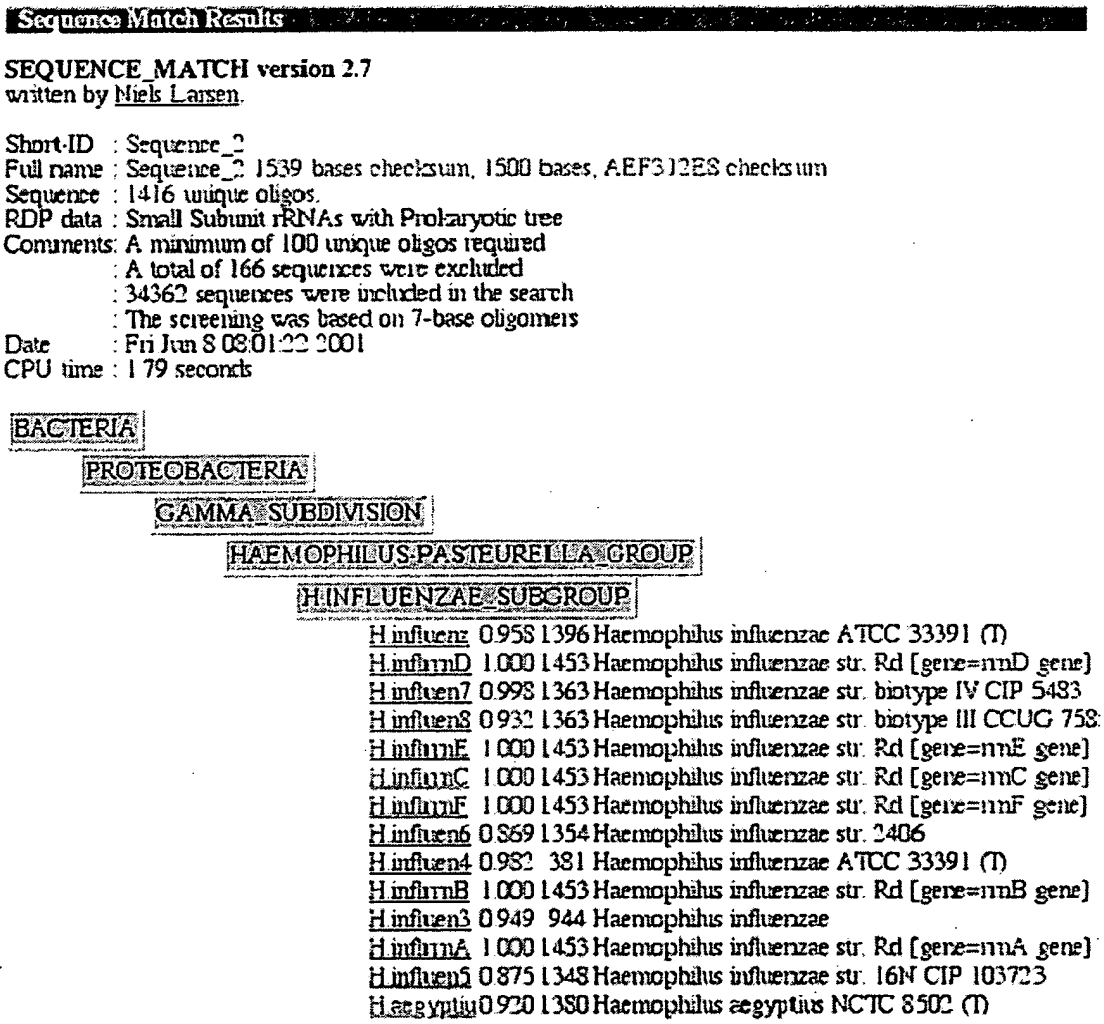
Note that the main browser window has advanced to the *Sequence Match* work page. The first step in the analysis is to load you sequence, As with the *Chimera Check* function described above and all similar analyses functions requiring user provided sequence, this can be accomplished by cutting and pasting or uploading a file. Clicking on one of the buttons below will automatically load one of the test sequences into the *Sequence Match* work page. Click now on Sequence 2.


Once you see the sequence loaded into the box on the browser window you are ready to submit the sequence to the RDP by clicking the labeled button at the bottom of the page. Note that there are several options in the Sequence Match Options section. We will return to these below. For now, click the ***Submit Sequences*** button, wait patiently for several seconds, and view the results.

## 2.3 Results from Sequence Match

If you had selected Sequence 2 for your first analysis and kept the default options, you will see the results presented in an HTML formatted phylogenetic arrangement (Figure 2.1). Note that the results begin with the Domain, in this case Bacteria, and proceeds through the phylogenetic scheme to the sequence(s) most closely related to the submitted sequence. The matching sequences are shown first by their short ID followed by the complete name. In between the short ID and the full name is a similarity estimate (S_ab) for the listed and submitted sequences. In the case of Sequence 2, note that there are several sequences identified with a perfect similarity score of 1.0 indicating that the sequences are identical. These results are not the consequence of multiple submissions of the same gene from the same organism or even different with the same rRNA sequence, but rather are derived from seven sequenced rRNA loci within a single organism. Sequence 2 was taken from the RDP archives and was derived from *Haemophilus influenzae*, the first bacteria to have its genome completely sequenced (1). Note that there are six perfect scores of 1.0, all derived from the same *H. influenzae* genome. These are indicated in the database as H.inflrrnA through H.inflrrnF, indicating that six ribosomal operons were found in the *H. influenzae* genome (for a more detailed documentation of rRNA and gene copy numbers, see the Ribosomal RNA Operon Copy Number Database (rrndb) at http://rrndb.cme.msu.edu.). Clicking on the identifier **H.inflrrnA** opens the annotated sequence file into the browser screen for inspection, providing the literature reference along with the aligned sequence. Close inspection of the similarity scores reveals that several scores within the Genus-species *H. influenzae* are

less than 95-96% (0.95-0.96) similar, suggesting that these organisms may be different at the species level from the submitted sequence. In order to confirm this, a student of phylogeny would need to determine the similarity of the complete sequences from these two strains as a starting point. At this point in our analyses, we do not know if we are looking at the results from comparisons of complete sequences. Note also that a list of unaligned sequences that closely match the submission follows the phylogenetic presentation. These sequences are from the unaligned collection of the RDP.



Sequence Match Results

**SEQUENCE_MATCH version 2.7**
written by Niels Larsen.

Short-ID : Sequence_2
Full name : Sequence_2 1539 bases checksum, 1500 bases, AEF3D2E8 checksum
Sequence : 1416 unique oligos.
RDP data : Small Subunit rRNAs with Prokaryotic tree
Comments: A minimum of 100 unique oligos required
: A total of 166 sequences were excluded
: 34362 sequences were included in the search
: The screening was based on 7-base oligomers
Date : Fri Jun 8 08:01:22 2001
CPU time : 1.79 seconds

BACTERIA
PROTEOBACTERIA
GAMMA SUBDIVISION
HAEMOPHILUS-PASTEURELLA GROUP
H.INFLUENZAE SUBGROUP

| | | | |
|---|---|---|---|
| H.influenz | 0.958 | 1396 | Haemophilus influenzae ATCC 33391 (T) |
| H.infrmD | 1.000 | 1453 | Haemophilus influenzae str. Rd [gene=rmD gene] |
| H.influen7 | 0.998 | 1363 | Haemophilus influenzae str. biotype IV CIP 5483 |
| H.influen8 | 0.932 | 1363 | Haemophilus influenzae str. biotype III CCUG 758 |
| H.infrmE | 1.000 | 1453 | Haemophilus influenzae str. Rd [gene=rmE gene] |
| H.infrmC | 1.000 | 1453 | Haemophilus influenzae str. Rd [gene=rmC gene] |
| H.infrmF | 1.000 | 1453 | Haemophilus influenzae str. Rd [gene=rmF gene] |
| H.influen6 | 0.869 | 1354 | Haemophilus influenzae str. 2406 |
| H.influen4 | 0.982 | 381 | Haemophilus influenzae ATCC 33391 (T) |
| H.infrmB | 1.000 | 1453 | Haemophilus influenzae str. Rd [gene=rmB gene] |
| H.influen3 | 0.949 | 944 | Haemophilus influenzae |
| H.infrmA | 1.000 | 1453 | Haemophilus influenzae str. Rd [gene=rmA gene] |
| H.influen5 | 0.875 | 1348 | Haemophilus influenzae str. 16N CIP 103723 |
| H.aegyptiu | 0.920 | 1380 | Haemophilus aegyptius NCTC 8502 (T) |

Figure 2.1. Results of Sequence Match with Sequence_2.

## 2.4 How Sequence Match Works

As detailed in the *Sequence Match* information pages, the match function is performed by first establishing a collection of unique oligomers seven bases long derived from the submitted sequence. Each oligomer is then compared to an oligomer dictionary derived from the RDP database and the degree of sequence identity is scored and saved. The sum of all the oligomer comparisons is then used to determine the sequences from the database with the greatest similarity to the submitted sequence. The information at the beginning of the reported results for **Sequence_2** indicates that 1,416 unique oligomers were extracted from the submitted sequence and that 34,362 sequences in the RDP were

compared to the submitted sequence. The *Sequence Match* function is not dependent on an aligned sequence database, hence both the aligned and unaligned databases are queried. In fact all sequences, bacterial, eukaryotic and mitochondrial in unaligned format are queried.

The reader will find that Sequence_3, when run through the *Sequence Match* function, matches to only three sequences in the aligned database from the ß subdivision of the Proteobacteria, as of 6/10/2001. Of all the RDP matches to Sequence 3, there are none that are even close to 95% similarity, indicating that this sequence (and organism) has not been seen before. Sequence 4 has a single perfect match within the RDP.

### 2.5 Sequence Match Options

- The first option requires the user to select the phylogenetic tree (prokaryotic, eukaryotic or mitochondrial) for display (see Figure 2.2 below). If the user has no inkling which to pick, try one to determine what sequence scores derived from the unaligned section are closest to the submitted sequence, then resubmit based on this affiliation.
- The second option is in the event that the user is working with the complementary strand. Recall that the RDP stores the RNA sequences. If your submitted sequences hit little or nothing, try it with the *Complement* selected.
- The third and fourth options allow the user to specify the inclusion of partial sequences from the database and the number of matching sequences to include in the output.
- The format option permits the user to specify between ASCI or HTML in TREE or LIST formats.
- The use is provided with the option of viewing the loaded sequence on the results page
- Lastly there is the option of having the results emailed.

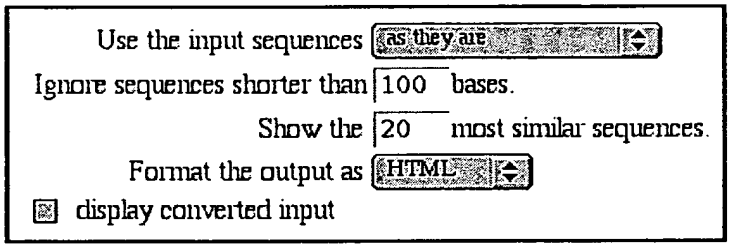
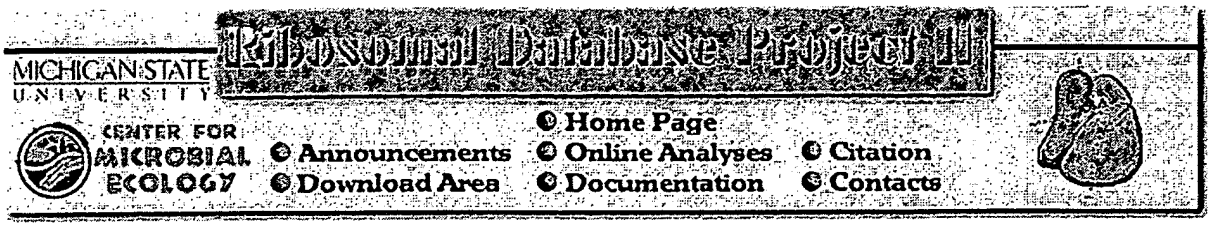



Figure 2.2




## About Sequence Match

Description

---

## Sequence Match Results - Example

**SEQUENCE_MATCH version 2.7**
written by Niels Larsen.

Short-ID :raw_1637
Full name :raw_1637 [Unknown form], 1514 bases, 4B9107AD checksum.
Sequence :1437 unique oligos.
RDP data :Small Subunit rRNAs with Prokaryotic tree
Comments:A minimum of 100 unique oligos required
:A total of 166 sequences were excluded
:21306 sequences were included in the search
:The screening was based on 7-base oligomers
Date :Mon Sep 13 14:57:11 1999
CPU time :1.99 seconds

BACTERIA
GRAM_POSITIVE_BACTERIA
SPOROMUSA_AND_RELATIVES
HELIOBACTERIUM_GROUP
DFI.DEHALOGENANS_SUBGROUP
env.A50u.590349clone A50u
CLOSTRIDIUM_AND_RELATIVES
C.LEPTUM_GROUP
STR.16SX_SUBGROUP
str.16SX-2 .9871433str. 16SX-2
str.16SX-11.0001437str. 16SX-1

UNCLASSIFIED / UNALIGNED
UEU68617.633515Unidentified eubacterium from the Amazon 16S ribosomal RNA gene,
UEU68614.629526Unidentified eubacterium from the Amazon 16S ribosomal RNA gene,
SAU11787.614101Staphylococcus aureus methicillin-resistant ATCC 33952 clone RRNV43
SAU11789.614101Staphylococcus aureus methicillin-resistant isolate H11 clone RRNV8
SAU11783.614101Staphylococcus aureus methicillin-resistant isolate H11 clone RRNV4
SAU11786.614101Staphylococcus aureus methicillin-resistant ATCC 33952 clone RRNV42

SAU11773.614101Staphylococcus aureus methicillin-resistant isolate D46 clone RRN4
SAU11780.614101Staphylococcus aureus methicillin-resistant ATCC 33952 clone RRNV32
SAU11779.614101Staphylococcus aureus methicillin-resistant ATCC 33952 clone RRNV30
SAU11782.614101Staphylococcus aureus methicillin-resistant ATCC 33952 clone RRNV38
SAU11778.614101Staphylococcus aureus methicillin-resistant ATCC 33952 clone RRNV27
SAU11776.614101Staphylococcus aureus methicillin-resistant isolate H11 clone
SAU11775.614101Staphylococcus aureus methicillin-resistant isolate H11 clone
SAU11784.614101Staphylococcus aureus methicillin-resistant ATCC 33952 clone RRNV40
SAU11781.604101Staphylococcus aureus methicillin-resistant ATCC 33952 clone RRNV34
UEU68628.596493Unidentified eubacterium from the Amazon 16S ribosomal RNA gene,
UEU68629.588488Unidentified eubacterium from the Amazon 16S ribosomal RNA gene,

---

## Description

This is a crude but fast way of measuring similarity between your sequence(s) and those in RDP. The example output shows the sequences that are most similar to the user sequence, which in this example is GenBank accession U27710. If there were several entries in the user file, there would be lists like this one for each entry.

The RDP maintained taxonomic list is used for presentation. This taxonomic list is extracted from the phylogenetic tree that RDP maintains, which assumes an alignment. But not all sequences are aligned (RDP is not up to date), so the sequences pending alignment are listed separately, as the group at the bottom, and sorted so the best match is at the top.

Each match line contains four elements, from left to right,

- A short ID used to uniquely identify the RDP sequence. A click will return the simple entry, including the sequence.
- A color coded similarity score (S_ab). These are the number of (unique) oligomers shared between your sequence and a given RDP sequence divided by the lowest number of unique oligos in either of the two sequences.
- The number of uniquely occurring oligomers within a given sequence (Olis). If the same oligomer occurs more than once then they are counted only once; thus this number only approximately reflects the sequence length. Counting only unique oligos compensates somewhat for composition bias (for example, inserts tend to be GC-rich and it becomes very likely that the same GC-rich oligos occur several times; by counting these only once, this artifact becomes less severe).
- Full name. The definition line from the RDP distribution, often the same as Genus/species/string name.

### What Happens to Your Data

After your sequences have been fetched by the RDP server from your local machine to a scratch area (from which it is always automatically deleted), the following happens:

- Each 7-oligomer (8-mers for 23S) in your sequence is translated into a unique integer (ranging from zero to 4 to the power of 7, minus one.) Only Watson-Crick base symbols (A, U/T, G and C) are used; all other characters are ignored.
- An oligomer dictionary is created with all entries flagged that are represented one or more times in your sequence.
- Pre-compiled match information is loaded: a long one-dimensional table with zero-terminated

runs of integers each of which corresponds to a given RDP sequence; the first entry in the distributed GenBank file would be 1, the 337'th would be 337 and so on. Then, an oligomer dictionary is loaded for the server dataset; it is similar to the one for your sequence, except the element values are not flags but integers that point to the start of each zero-terminated run. Example,

```
Oligomer dictionary:  |                    |              |   |      etc
Table of seq numbers: 001 004 013 246 000  001 007 056 000 000 344 000 etc
```

- A one-dimensional integer array is created for counting matches, as a results table. It has the same number of elements as the number of organisms in the server dataset. This result array will contain the number of oligos shared between a given server sequence and yours.
- Move sequentially through your sequence. The oligomer starting at each position is looked up in the RDP oligo dictionary. The number returned by this dictionary is the starting point of the corresponding integer run, which shows which sequences have a given oligo. Then read all integers until a zero is encountered, while incrementing the corresponding slots in the results table. Note that the comparison does not involve any RDP sequences, just pre-compiled information of which sequences a given oligo occurs in; thus the speed.
- Calculate similarity score from the results array: Each element is divided by the number of unique oligomers (ie if two or more occurrences of same, they are counted as one) in either your sequence or the server sequence that it is being compared with, whichever number is lower. This simple way, the quality of the match becomes independent of the length of the sequences.
- Sort the similarity scores in descending order.
- Display as many of the matches as the user specifies.

**Other output options**

The HTML list is a flat list where the best matches are ranked at the top, example below. There are also ascii versions of both the taxonomic and flat lists, good for electronic mail.

---

**SEQUENCE_MATCH version 2.7**

written by Niels Larsen

Short-ID :Mc.maripal
Full name :Methanococcus maripaludis str. JJ.
Sequence :1293 unique oligos.
RDP data :Small Subunit rRNAs
Comments:A minimum of 100 unique oligos required
:A total of 16 sequences were excluded
:10707 sequences were included in the search
:The screening was based on 7-base oligomers
Date :Thu Jan 29 18:55:07 1998
CPU time :.46 seconds

| Short-ID | S_ab | Olis | Full name |
|---|---|---|---|
| Mc.maripal | -1 | 293 | Methanococcus maripaludis str. JJ. |

| | | | | |
|---|---|---|---|---|
| Mc.maripal | 1.000 | 1293 | Methanococcus maripaludis str. JJ. |
| Mc.maripa2 | 1.000 | 1293 | Methanococcus maripaludis str. JJ. |
| Mc.deltae | .903 | 1306 | Methanococcus deltae str. D. |
| Mc.deltae2 | .903 | 1303 | Methanococcus deltae str. delta RC. |
| Mc.maripa3 | .894 | 1317 | Methanococcus maripaludis str. C5. |
| Mc.maripa4 | .886 | 1302 | Methanococcus maripaludis str. C6. |
| Mc.maripa5 | .862 | 1300 | Methanococcus maripaludis str. C7. |
| Mc.vanniel | .757 | 1385 | Methanococcus vannielii str. EY33. |
| Mc.voltae3 | .676 | 1321 | Methanococcus voltae str. A3. |
| Mc.voltae2 | .670 | 1353 | Methanococcus voltae str. PS. |
| Mc.voltae | .670 | 1353 | Methanococcus voltae str. PS. |
| Mc.aeolicu | .643 | 1319 | "Methanococcus aeolicus" str. A. |
| Mc.aeolic2 | .643 | 1319 | "Methanococcus aeolicus". |
| Mc.thlitho | .643 | 1352 | Methanococcus thermolithotrophicus str. SN-1. |
| Mc.igneus | .500 | 1223 | Methanococcus igneus str. Kol 5. |
| Mc.jannasc | .445 | 1298 | Methanococcus jannaschii str. JAL-1. |
| Tc.peptphl | .399 | 873 | Thermococcus sp. str. OG-1. |
| THC16SRD1 | .399 | 873 | Thermococcus peptonophilus gene for 16S ribosomal RNA, partial |
| Tc.celer2 | .387 | 256 | Thermococcus celer. |
| Pc.furiosu | .385 | 200 | Pyrococcus furiosus. |

Finally, there is an easy-to-parse tabular format, for those who want to extract some of the result information into a spreadsheet etc. This format will not change as often as the others. Here is an example,

```
HEAD> Program SEQUENCE_MATCH
HEAD> Version 2.7
HEAD> Author Niels Larsen
HEAD> Email niels@cme.msu.edu
HEAD> ShortID Mc.maripal
HEAD> Fullname Methanococcus maripaludis  str. JJ.
HEAD> Olinum 1293
HEAD> Complemented no
HEAD> RDPdata SSU_Unal_simrank.bin
HEAD> Minseq 100
HEAD> Exclude 16
HEAD> Include 10707
HEAD> Olilen 7
HEAD> Date Thu Jan 29 19:43:14 1998
HEAD> CPU .48 seconds
HEAD> Outwidth 78
TABLE> Mc.maripal 1.000 1293 Methanococcu s maripaludis str. JJ.
TABLE> Mc.maripa2 1.000 1293 Methanococcu s maripaludis str. JJ.
TABLE> Mc.deltae .903 1306 Methanococcus  deltae str. D.
TABLE> Mc.deltae2 .903 1303 Methanococcus  deltae str. delta RC.
TABLE> Mc.maripa3 .894 1317 Methanococcus  maripaludis str. C5.
TABLE> Mc.maripa4 .886 1302 Methanococcus  maripaludis str. C6.
TABLE> Mc.maripa5 .862 1300 Methanococcus  maripaludis str. C7.
TABLE> Mc.vanniel .757 1385 Methanococcus  vannielii str. EY33.
TABLE> Mc.voltae3 .676 1321 Methanococcus  voltae str. A3.
TABLE> Mc.voltae2 .670 1353 Methanococcus  voltae str. PS.
TABLE> Mc.voltae .670 1353 Methanococcus  voltae str. PS.
```

```
TABLE> Mc.aeolicu .643 1319 "Methanococcu s aeolicus" str. A.
TABLE> Mc.aeolic2 .643 1319 "Methanococcu s aeolicus".
TABLE> Mc.thlitho .643 1352 Methanococcus  thermolithotrophicus str. SN -1.
TABLE> Mc.igneus .500 1223 Methanococcus  igneus str. Kol 5.
TABLE> Mc.jannasc .445 1298 Methanococcus  jannaschii str. JAL -1.
TABLE> Tc.peptphl .399 873 Thermococcus s p. str. OG-1.
TABLE> THC16SRD1 .399 873 Thermococcus peptonophilus gene for 16S ribosomal RNA
TABLE> Tc.celer2 .387 256 Thermococcus ce ler.
TABLE> Pc.furiosu .385 200 Pyrococcus fur iosus.
```

The HEAD> section contains tab-separated key/value pairs, and the TABLE> section contains four tab-separated values.

The following Perl code parses the previous tabular output:

```
#!/usr/bin/perl
# The following Perl code fills the above  Sequence Match tab- table
# output into a header and table hash tha t both use the Short -ID as
# keys.
$line = <STDIN>;
while ($line =~ /\w/) {
    undef %header;
    while ($line =~ /^HEAD>\t([^\t]*)\t(. *)$/) {
        $header{$1} = $2;
        $line = <STDIN>;
    }
    foreach $key (keys %header) {
        $headers{$header{"ShortID"}}{$key } = $header{$key};
    }
    $i = 0;
    while ($line =~ /^TABLE>\t([^\t]*)\t( [^\t]*)\t([^\t]*)\t(.*)$/) {
        $tables{$header{"ShortID"}}[$i++]  =
        {"ShortID" => $1,
         "Sab" => $2,
         "Olis" => $3,
         "Fullname" => $4};
        $line = <STDIN>;
    }
}
# The output is now easily accessible in  separate header and table
# hashes. For example, if you want the Fu ll name for Mc.maripal,
#
#   $headers{"Mc.maripal"}{"Fullname"};
#
# or if you want the match values for the  best and second best
# match (top two lines) with Mc.maripal,
#
#   $tables{"Mc.maripal"}[0]{"Sab"};
#   $tables{"Mc.maripal"}[1]{"Sab"};
#
```

---

*Questions? Mail them to RDP-II Web Support.*

**EXHIBIT E**

(See e.g., Exhibit E Maidak et al., *Nucl. Acid. Res.* 24:82-85 (1996))

82–85 Nucleic Acids Research, 1996, Vol. 24, No. 1 

# The Ribosomal Database Project (RDP)

**Bonnie L. Maidak*, Gary J. Olsen, Niels Larsen[1], Ross Overbeek[2], Michael J. McCaughey and Carl R. Woese**

Department of Microbiology, University of Illinois, 131 Burrill Hall, 407 South Goodwin Avenue, Urbana, IL 61801-3704, USA, [1]Center for Microbial Ecology, Michigan State University, East Lansing, MI 48824, USA and [2]Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439, USA

Received October 4, 1995; Accepted October 5, 1995

## ABSTRACT

**The Ribosomal Database Project (RDP) is a curated database that offers ribosome-related data, analysis services and associated computer programs. The offerings include phylogenetically ordered alignments of ribosomal RNA (rRNA) sequences, derived phylogenetic trees, rRNA secondary structure diagrams and various software for handling, analyzing and displaying alignments and trees. The data are available via anonymous ftp (rdp.life.uiuc.edu), electronic mail (server@rdp.life.uiuc.edu), gopher (rdpgopher.life.uiuc.edu) and World Wide Web (WWW) (http://rdpwww.life.uiuc.edu/). The electronic mail and WWW servers provide ribosomal probe checking, screening for possible chimeric rRNA sequences, automated alignment and approximate phylogenetic placement of user-submitted sequences on an existing phylogenetic tree.**

## DESCRIPTION

The Ribosomal Database Project (RDP) provides data, programs and services related to the ribosome. In this paper we summarize these offerings, the changes that have been introduced since last year's description (1) and some future features.

### Data

The ribosomal RNA sequences in the RDP alignments are drawn from major sequence repositories [GenBank (2) and EBI (3)] and direct submissions to the RDP. They are organized and presented in an aligned and phylogenetically ordered form. Each sequence is annotated with its organismal source (for cultured organisms: the genus, species, culture collection numbers, etc.), cellular compartment, origin of sequence data (usually a literature citation) and other relevant information. If multiple versions of a given sequence exist, the RDP attempts to select by a variety of criteria (which include the frequency of putative sequence errors and completeness) only one of the versions for release. As a consequence, the number of released sequences is lower than the number of existing sequences. The RDP staff also examines the original publications and updates annotations, strain designations and organism names. Submitters and/or the public sequence databases are notified of possible errors.

The small subunit (SSU) rRNA alignments currently comprise sequences from ~140 Archaea, 2700 Bacteria (including chloroplasts and a few plant mitochondria) and 440 Eucarya (an alignment supplied by M. L. Sogin, Woods Hole Marine Biology Laboratory, MA). A representative alignment of 98 prokaryotic small subunit rRNA sequences is also available. The number of large subunit (LSU) rRNA sequences remains at 150.

A phylogenetic tree is available for the sequences in the posted prokaryotic and eukaryotic (new this year) SSU rRNA alignments. They have been assembled from appropriately overlapping subtrees, each of which has been inferred using maximum-likelihood analysis (4,5). The current trees (and subsets of them) are available in printable text, PostScript and Newick formats. The RDP also offers a collection of SSU and LSU rRNA secondary structure diagrams in PostScript format generated and supplied by R. Gutell *et al.* (6). Also new this year is the corresponding taxonomic listing for SSU eukaryotic rRNA sequences.

As stated in last year's article, the RDP has made available this year unaligned data sets for use with the SIMILARITY_RANK, CHECK_PROBE and CHECK_CHIMERA commands (see description of these commands in Table 1). The current release contains 5500 and 1400 sequences respectively for the SSU and LSU data sets.

To facilitate access to specific rRNA aligned and unaligned sequences, the RDP offers subdirectories containing GenBank-formatted files of each sequence (directory names: aligned/sequences/[A-Z] and unaligned/sequences/[A-Z]).

### Data servers

During the past year, an RDP WWW server has been developed. The initial part of its Home Page is shown in Figure 1.

Table 1 lists the commands available on the electronic mail server, which are also available on the WWW server.

### Programs

The programs currently available through the RDP servers are listed in Table 2.

* To whom correspondence should be addressed

**Figure 1.** The RDP home page

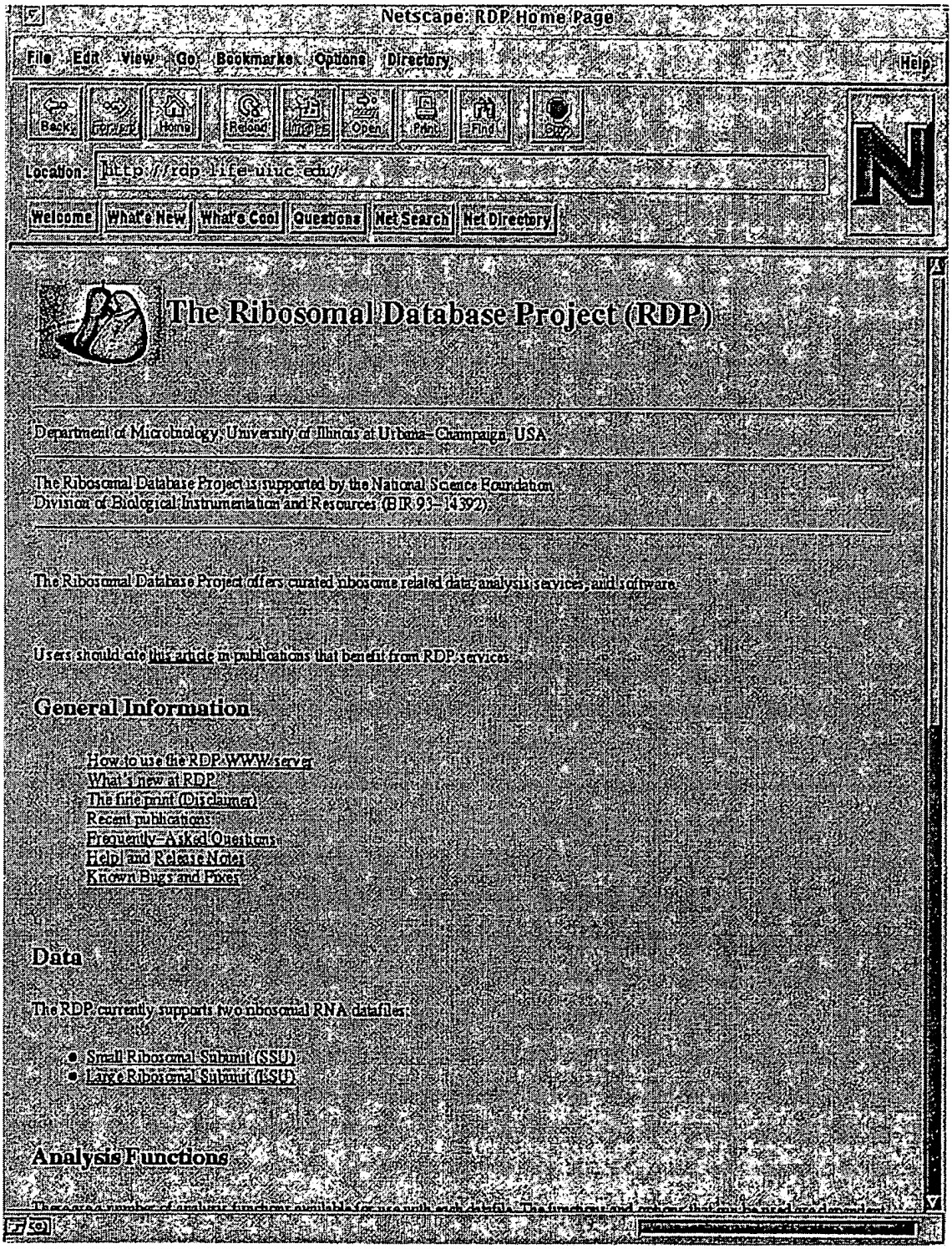


**Table 1.** Electronic mail server commands

| | |
|---|---|
| *General functions* | |
| HELP | Obtain general instructions for using the RDP mail server, or obtain a detailed description of a specified command. |
| SUBSCRIBE (UNSUBSCRIBE) | Have your name put on (or taken off) the RDP electronic mailing list for notifications about new data and services. |
| *Directory and file functions* | |
| DIRECTORY | Obtain a listing of the files in an RDP directory or directory hierarchy. |
| INFORMATION | Obtain a description of the data in a specified RDP directory. |
| GET | Obtain a copy of a specified file. |
| *Data retrieval* | |
| FULL_ALIGNMENT | Obtain a copy of a complete sequence alignment. Options allow selection of the format. |
| SUBALIGNMENT | Obtain a subalignment containing specified sequences and/or positions from a larger alignment. Options allow selection of the format. |
| FULL_TREE | Obtain a copy of a phylogenetic tree in a requested format (printable text, PostScript or Newick). |
| SUBTREE | Obtain a tree containing specified sequences from a larger tree in a requested format. |
| NAMES | Obtain a list of the names of the sequences represented in a specified alignment or tree. |
| *Analytic functions* | |
| ALIGN_SEQUENCE | Align a user-supplied sequence on the most similar sequence from the RDP. An option allows the user to avoid short matches with partial sequences. |
| CHECK_CHIMERA | Analyze a user-supplied sequence for evidence of chimeric structure. Options allow the user to add their own sequences to the database used in the analysis and to ignore short matches with partial sequences. Detects possible chimeric sequences |
| CHECK_PROBE | Analyze the occurrences of a specified 'probe' sequence in a set of sequences. |
| SIMILARITY_RANK | Obtain a list of the sequences most similar to that submitted. Options allow the user to add their own sequences to the database used in the analysis and to ignore short matches with partial sequences. |
| SUGGEST_TREE | Obtain an approximate placement in the RDP tree of a user-submitted sequence (using maximum likelihood analysis). |
| *Defining the data to be used in analyses* | |
| RDP_LIST | Use all available data in subsequent server commands. |
| REP_LIST | Use a standard representative subset of the available data in subsequent server commands. |
| MY_DATABASE | Add the user-provided sequences to the database used in the commands SIMILARITY_RANK and CHECK_CHIMERA. |
| MY_LIST | Use the specified subset of available data in subsequent server commands. |
| MY_SEQUENCES | Provide sequence data for use in subsequent server commands. |

Mail messages utilizing these commands should be sent to server@rdp.life.uiuc.edu.

**Table 2.** Programs available through the RDP servers

| | |
|---|---|
| Convert_aln | A sequence alignment format conversion program for UNIX and VAX/VMS systems. |
| DNArates | A maximum likelihood method to estimate site-specific rates of nucleotide substitution from a sequence alignment and a user-defined phylogenetic tree. Data formats are similar to those used in J. Felsenstein's PHYLIP package. Compatible with a wide variety of computers. |
| Editor_AE2 | An alignment editor and analysis program written by T. Macke for UNIX systems. |
| Editor_GDE | The Genetic Data Environment sequence alignment editing and analysis package written by S. Smith. Posted version is for Sun Microsystems computers. |
| EPSFilter | Macintosh program for working with Encapsulated PostScript (EPS) files written by B. Fowler. |
| fastDNAml | A maximum likelihood tree inference program based on version 3.3 of J. Felsenstein's DNAML. It has features to facilitate analysis of a larger number of taxa. Compatible with a wide variety of computers. |
| GraphicConverter | Macintosh program for conversion between graphics formats written by T. Lemke. |
| Readseq | A suite of sequence format conversion programs written by D. Gilbert. Compatible with a wide variety of computers. |
| SeqEdit | An alignment editor and analysis program for VAX/VMS systems. |
| Subalign | A program to extract specified rows and columns from an alignment. For UNIX and VAX/VMS systems. |
| TreeTool | A X-windows-based phylogenetic tree manipulation program for Sun Microsystems computers |

## RDP CITATION AND ACCESS

Research assisted by any RDP service should cite: the Ribosomal Database Project (RDP) at the University of Illinois in Urbana, IL; the release number and this article (i.e. Maidak *et al.*, 1996). Please state which data, programs and services were used and the method of access.

The RDP data and analysis services can be found at URL: http://rdpwww.life.uiuc.edu/.

The RDP data can be accessed via anonymous ftp to rdp.life.uiuc.edu. Once you are logged in (using a user-id of 'anonymous' and your electronic mail address for password), examine the 00README files, which describe the organization of the data and programs.

The address of the automated electronic mail server is server@rdp.life.uiuc.edu. To obtain an overview of what data and services are currently available, send a mail message with the phrase 'help' as the body of the message. (Full command descriptions can be obtained by sending 'help complete'). If your electronic mail address is unknown to the e-mail server, you will also receive a registration form. After returning the completed registration form, you will be automatically notified when new data or services become available.

The RDP gopher host name is rdpgopher.life.uiuc.edu. Gopher access to RDP data through the WWW is also available (URL: gopher://rdpgopher.life.uiuc.edu/).

Electronic mail correspondence with RDP staff should be addressed to rdp@phylo.life.uiuc.edu. Those without access to electronic mail may contact the RDP curator (B.L.M.) via telephone (+1 217 333 5866), FAX (+1 217 244 6697) or regular mail.

## FUTURE CHANGES AND ADDITIONS

Future plans for the RDP include improvements (i) in the display of and interaction with the phylogenetic trees, (ii) in the presentation and output options of the SUGGEST_TREE command and (iii) an improved version of the CHECK_PROBE function. Also planned is a sequence evaluation program, which assesses the quality of a user-supplied sequence, reporting back possible sequencing errors and/or idiosyncrasies, as well as a 'sequence signature' which defines the approximate taxonomic position of the sequence.

## ACKNOWLEDGEMENTS

We thank R. Gutell (and his colleagues) and M. L. Sogin for providing their data collections. The RDP is largely supported by the National Science Foundation, Biological Instrumentation and Resources Division.

## REFERENCES

1 Maidak,B.L., Larsen,N., McCaughey,M.J., Overbeek,R., Olsen,G.J., Fogel,K., Blandy,J. and Woese,C.R. (1994) *Nucleic Acids Res.*,**22**, 3485–3487.
2 Benson,D.A., Boguski,M., Lipman,D.J. and Ostell,J. (1994) *Nucleic Acids Res.*, **22**, 3441–3444.
3 Emmert,D.B., Stoehr,P.J., Stoesser,G. and Cameron,G.N. (1994) *Nucleic Acids Res.*,**22**, 3445–3449
4 Felsenstein,J. (1981) *J. Mol. Evol.*,**17**, 368–376.
5 Olsen,G.J., Matsuda,H., Hagstrom,R. and Overbeek,R. (1994) *CABIOS*, **10**, 41–48.
6 Gutell,R.R. *et al.* (1996) *Nucleic Acids Res.*,**24**, this issue.

**EXHIBIT F**

Exemplary partial 16s rRNA sequences of isolated *Salinospora* strains of the present invention.

Signature nucleotides are bolded.

AY040623 . Salinospora sp. C...[gi:22474397]
Links

```
LOCUS       AY040623                1480 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH964 16S ribosomal RNA gene, partial sequence.
ACCESSION   AY040623
VERSION     AY040623.1  GI:22474397
KEYWORDS    .
SOURCE      Salinospora sp. CNH964
  ORGANISM Salinospora sp. CNH964 Bacteria; Actinobacteria; Actinobacteridae;
Actinomycetales;
            Micromonosporineae; Micromonosporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent populations of a major new marine
            actinomycete taxon in ocean sediments
  JOURNAL   Appl. Environ. Microbiol. 68 (10), 5005-5011 (2002)
  MEDLINE 22235406 PUBMED 12324350 REFERENCE   2  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Institution of
            Oceanography, UCSD, 8602 La Jolla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1480
                     /organism="Salinospora sp. CNH964"
                     /mol_type="genomic DNA"
                     /strain="CNH964"
                     /db_xref="taxon:168701" rRNA <1..>1480
                     /product="16S ribosomal RNA"
ORIGIN
        1 agagtttgat cctggctcag gacgaacgct ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggcg aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaaaccgg ggctaatacc ggatatgacc atctgtcgca
      181 tggtgggtgg tggaaagatt ttttggcttg ggatgggctc gcggcctatc agcttgttgg
      241 tggggtgatg gcctaccaag gcggcgacgg gtagccggcc tgagagggcg accggccaca
      301 ctgggactga gacacggccc agactcctac gggaggcagc agtggggaat cttgcacaat
      361 gggcggaagc ctgatgcagc gacgccgcgt gagggatgac ggccttcggg ttgtaaacct
      421 ctttcagcag ggacgaagcg ttttgtgacgg tacctgcaga agaagcgccg gccaactacg
      481 tgccagcagc cgcggtaaga cgtagggcgc aagcgttgtc cggatttatt gggcgtaaag
      541 agctcgtagg cggcttgtcg cgtcgactgt gaaaacctgt ggctcaactg cgggcttgca
      601 gtcgatacgg gcaggctaga gttcggtagg ggagactgga attcctggtg tagcggtgaa
      661 atgcgcagat atcaggagga acaccggtgg cgaaggcggg tctctgggcc gatactgacg
      721 ctgaggagcg aaagcgtggg gagcgaacag gattagatac cctggtagtc cacgctgtaa
      781 acgttgggcg ctaggtgtgg ggggcctctc cggttctctg tgccgcagct aacgcattaa
      841 gcgccccgcc tggggagtac ggccgcaagg ctaaaactca aaggaattga cgggggcccg
      901 cacaagcggc ggagcatgcg gattaattcg atgcaacgcg aagaacctta cctgggtttg
      961 acatcgccgg aaatccttca gagatggggg gtccttcggg gccggtgaca ggtggtgcat
     1021 ggctgtcgtc agctcgtgtc gtgagatgtt gggttaagtc ccgcaacgag cgcaaccctt
     1081 gttcgatgtt gccagcgcgt tatggcgggg actcatcgaa gactgccggg gtcaactcgg
     1141 aggaaggtgg ggatgacgtc aagtcatcat gccccttatg tccagggctt cacgcatgct
     1201 acaatggccg gtacagtggg ctgcgatacc gtgaggtgga gcgaatccca aaaagccggt
     1261 ctcagttcgg atcggggtct gcaactcgac cccgtgaagt cggagtcgct agtaatcgca
     1321 gataaacaac gctgcggtga atacgttccc gggccttgta cacaccgccc gtcacgtcac
     1381 gaaagtcggc aacacccgaa gccggtggcc taacccttgt ggggggagcc gtcgaaggtg
     1441 gggctggcga ttgggacgaa gtcgtaacaa ggtagccgta
//
```

AY040622 . Salinospora sp. C...[gi:22474395]
Links

```
LOCUS       AY040622                1480 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH898 16S ribosomal RNA gene, partial sequence.
ACCESSION   AY040622
VERSION     AY040622.1  GI:22474395
KEYWORDS    .
SOURCE      Salinospora sp. CNH898
  ORGANISM Salinospora sp. CNH898 Bacteria; Actinobacteria; Actinobacteridae;
Actinomycetales;
            Micromonosporineae; Micromonosporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent populations of a major new marine
```

```
            actinomycete taxon in ocean sediments
  JOURNAL   Appl. Environ. Microbiol. 68 (10), 5005-5011 (2002)
  MEDLINE 22235406 PUBMED 12324350 REFERENCE   2  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Institution of
            Oceanography, UCSD, 8602 La Jolla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1480
                     /organism="Salinospora sp. CNH898"
                     /mol_type="genomic DNA"
                     /strain="CNH898"
                     /db_xref="taxon:168700" rRNA <1..>1480
                     /product="16S ribosomal RNA"
ORIGIN
        1 agagtttgat cctggctcag gacgaacgct ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggcg aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaaaccgg ggctaatacc ggatatgact ggctgccgca
      181 tggtggttgg tggaaagatt ttttggcttg ggatgggctc gcggcctatc agcttgttgg
      241 tggggtgatg gcctaccaag gcggcgacgg gtagccggcc tgagagggcg accggccaca
      301 ctgggactga gacacggccc agactcctac gggaggcagc agtggggaat cttgcacaat
      361 gggcggaagc ctgatgcagc gacgccgcgt gagggatgac ggccttcggg ttgtaaacct
      421 ctttcagcag ggacgaagcg ttttgtgacgg tacctgcaga agaagcgccg gccaactacg
      481 tgccagcagc cgcggtaaga cgtagggcgc aagcgttgtc cggatttatt gggcgtaaag
      541 agctcgtagg cggcttgtcg cgtcgactgt gaaaacccgt ggctcaactg cgggcttgca
      601 gtcgatacgg gcaggctaga gttcggtagg ggagactgga attcctggtg tagcggtgaa
      661 atgcgcagat atcaggagga acaccggtgg cgaaggcggg tctctgggcc gatactgacg
      721 ctgaggagcg aaagcgtggg gagcgaacag gattagatac cctggtagtc cacgctgtaa
      781 acgttgggcg ctaggtgtgg ggagcctctc cggttctctg tgccgcagct aacgcattaa
      841 gcgccccgcc tggggagtac ggccgcaagg ctaaaactca aaggaattga cgggggcccg
      901 cacaagcggc ggagcatgcg gattaattcg atgcaacgcg aagaacctta cctgggtttg
      961 acatcgccgg aaatccttca gagatggggg gtccttcggg gccggtgaca ggtggtgcat
     1021 ggctgtcgtc agctcgtgtc gtgagatgtt gggttaagtc ccgcaacgag cgcaaccctt
     1081 gttcgatgtt gccagcgcgt tatggcgggg actcatcgaa gactgccggg gtcaactcgg
     1141 aggaaggtgg ggatgacgtc aagtcatcat gccccttatg tccagggctt cacgcatgct
     1201 acaatggccg gtacaatggg ctgcgatacc gtgaggtgga gcgaatccca aaaagccggt
     1261 ctcagttcgg atcggggtct gcaactcgac cccgtgaagt cggagtcgct agtaatcgca
     1321 gatcaacaac gctgcgggga atacgttccc gggccttgta cacaccgccc gtcacgtcac
     1381 gaaagtcggc aacacccgaa gccggtggcc taacccttgt gggggggagcc gtcgaaggtg
     1441 gggctggcga ttgggacgaa gtcgtaacaa ggtagccgta
//
```

AY040621 . Salinospora sp. C...[gi:22474394]

Links

```
LOCUS       AY040621                1480 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH725 16S ribosomal RNA gene, partial sequence.
ACCESSION   AY040621
VERSION     AY040621.1  GI:22474394
KEYWORDS    .
SOURCE      Salinospora sp. CNH725
  ORGANISM Salinospora sp. CNH725 Bacteria; Actinobacteria; Actinobacteridae;
Actinomycetales;
            Micromonosporineae; Micromonosporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent populations of a major new marine
            actinomycete taxon in ocean sediments
  JOURNAL   Appl. Environ. Microbiol. 68 (10), 5005-5011 (2002)
  MEDLINE 22235406 PUBMED 12324350 REFERENCE   2  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Institution of
            Oceanography, UCSD, 8602 La Jolla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1480
                     /organism="Salinospora sp. CNH725"
                     /mol_type="genomic DNA"
                     /strain="CNH725"
                     /db_xref="taxon:168699" rRNA <1..>1480
```

```
                     /product="16S ribosomal RNA"
ORIGIN
        1 agagtttgat cctggctcag gacgaacgct ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggcg aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaacccgg ggctaatacc ggatatgacc atctgtcgca
      181 tggtgggtgg tggaaagatt ttttggcttg ggatgggctc gcggcctatc agcttgttgg
      241 tggggtgatg gcctaccaag gcggcgacgg gtagccggcc tgagagggcg accggccaca
      301 ctgggactga gacacggccc agactcctac gggaggcagc agtggggaat cttgcacaat
      361 gggcggaagc ctgatgcagc gacgccgcgt gagggatgac ggccttcggg ttgtaaacct
      421 ctttcagcag ggacgaagcg ttttgtgacgg tacctgcaga aaaagcgccg gccaactacg
      481 tgccagcagc cgcggtaaga cgtagggcgc aagcgttgtc cggatttatt gggcgtaaag
      541 agctcgtagg cggcttgtcg cgtcgactgt gaaaacccgt ggctcaactg cgggcttgca
      601 gtcgatacgg gcaggctaga gttcggtagg ggagactgga attcctggtg tagcggtgaa
      661 atgcgcagat atcaggagga acaccggtgg cgaaggcggg tctctgggcc gatactgacg
      721 ctgaggagcg aaagcgtggg gagcgaacag gattagatac cctggtagtc cacgctgtaa
      781 acgttgggcg ctaggtgtgg ggggcctctc cggttctctg tgccgcagct aacgcattaa
      841 gcgccccgcc tggggagtac ggccgcaagg ctaaaactca aaggaattga cgggggcccg
      901 cacaagcggc ggagcatgcg gattaattcg atgcaacgcg aagaacctta cctgggtttg
      961 acatcgccgg aaatccttca gagatggggg gtccttcggg gccggtgaca ggtggtgcat
     1021 ggctgtcgtc agctcgtgtc gtgagatgtt gggttaagtc ccgcaacgag cgcaaccctt
     1081 gttcgatgtt gccagcgcgt tatggcgggg actcatcgaa gactgccggg gtcaactcgg
     1141 aggaaggtgg ggatgacgtc aagtcatcat gccccttatg tccagggctt cacgcatgct
     1201 acaatggccg gtacagtggg ctgcgatacc gtgaggtgga gcgaatccca aaaagccggt
     1261 ctcagttcgg atcggggtct gcaactcgac cccgtgaagt cggagtcgct agtaatcgca
     1321 gatcagcaac gctgcggtga atacgttccc gggccttgta cacaccgccc gtcacgtcac
     1381 gaaagtcggc aacacccgaa gccggtggcc taacccttgt ggggggagcc gtcaagtggg
     1441 gggctggcga ttgggacgaa gtcgtaacaa ggtagccgta
//

AY040620 . Salinospora sp. C...[gi:22474393]
Links
LOCUS       AY040620                1482 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH646 16S ribosomal RNA gene, partial sequence.
ACCESSION   AY040620
VERSION     AY040620.1  GI:22474393
KEYWORDS    .
SOURCE      Salinospora sp. CNH646
  ORGANISM Salinospora sp. CNH646 Bacteria; Actinobacteria; Actinobacteridae;
Actinomycetales;
            Micromonosporineae; Micromonosporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1482)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent populations of a major new marine
            actinomycete taxon in ocean sediments
  JOURNAL   Appl. Environ. Microbiol. 68 (10), 5005-5011 (2002)
  MEDLINE 22235406 PUBMED 12324350 REFERENCE   2  (bases 1 to 1482)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Institution of
            Oceanography, UCSD, 8602 La Jolla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1482
                     /organism="Salinospora sp. CNH646"
                     /mol_type="genomic DNA"
                     /strain="CNH646"
                     /db_xref="taxon:168698" rRNA <1..>1482
                     /product="16S ribosomal RNA"
ORIGIN
        1 agagtttgat cctggctcag gacgaacgct ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggcg aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaaaccgg ggctaatacc ggatatgacc atctgtcgca
      181 tggtgggtgg tggaaagatt ttttggcttg ggatgggctc gcggcctatc agcttgttgg
      241 tggggtgatg gcctaccaag gcggcgacgg gtagccggcc tgagagggcg accggccaca
      301 ctgggactga gacacggccc agactcctac gggaggcagc agtggggaat cttgcacaat
      361 gggcggaagc ctgatgcagc gacgccgcgt gagggatgac ggccttcggg ttgtaaacct
      421 ctttcagcag ggacgaagcg ttttgtgacgg tacctgcaga agaagcgccg gccaactacg
      481 tgccagcagc cgcggtaaga cgtagggcgc aagcgttgtc cggatttatt gggcgtaaag
      541 agctcgtagg cggcttgtcg cgtcgactgt gaaaacccgt ggctcaactg cgggcttgca
      601 gtcgatacgg gcaggctaga gttcggtagg ggagactgga attcctggtg tagcggtgaa
      661 atgcgcagat atcaggagga acaccggtgg cgaaggcggg tctctgggcc gatactgacg
```

```
      721 ctgaggagcg aaagcgtggg gagcgaacag gattagatac cctggtagtc cacgctgtaa
      781 acgttgggcg ctaggtgtgg ggggcctctc cggttctctg tgccgcagct aacgcattaa
      841 gcgccccgcc tggggagtac ggccgcaagg ctaaaactca aaggaattga cgggggcccg
      901 cacaagcggc ggagcatgcg gattaattcg atgcaacgcg aagaacctta cctgggtttg
      961 acatcgccgg aaatccttca gagatggggg gtccttcggg gccggtgaca ggtggtgcat
     1021 ggctgtcgtc agctcgtgtc gtgagatgtt gggttaagtc ccgcaacgag cgcaaccctt
     1081 gttcgatgtt gccagcgcgt tatggcgggg actcatcgaa gactgccggg gtcaactcgg
     1141 aggaaggtgg ggatgacgtc aagtcatcat gccccttatg tccagggctt cacgcatgct
     1201 acaatggccg gtacagtggg ctgcgatacc gtgaggtgga gcgaatccca aaaagccggt
     1261 ctcagttcgg atcggggtct gcaactcgac cccgtgaagt cggagtcgct agtaatcgca
     1321 gatcagcaac gctgcggtga atacgttccc gggccttgta cacaccgccc gtcacgtcac
     1381 gaaagtcggc aaccaacccg aagccggtgg cctaaccctt gtggggggag ccgtcgaagg
     1441 gggggctggc gattgggacg aagtcgtaac aaggtagccg ta
//

AY040619 . Salinospora sp. C...[gi:22474396]
Links
LOCUS       AY040619                1481 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH643 16S ribosomal RNA gene, partial sequence.
ACCESSION   AY040619
VERSION     AY040619.1  GI:22474396
KEYWORDS    .
SOURCE      Salinospora sp. CNH643
  ORGANISM Salinospora sp. CNH643 Bacteria; Actinobacteria; Actinobacteridae;
Actinomycetales;
            Micromonosporineae; Micromonosporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1481)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent populations of a major new marine
            actinomycete taxon in ocean sediments
  JOURNAL   Appl. Environ. Microbiol. 68 (10), 5005-5011 (2002)
  MEDLINE 22235406 PUBMED 12324350 REFERENCE   2  (bases 1 to 1481)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Institution of
            Oceanography, UCSD, 8602 La Jolla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1481
                     /organism="Salinospora sp. CNH643"
                     /mol_type="genomic DNA"
                     /strain="CNH643"
                     /db_xref="taxon:168697" rRNA <1..>1481
                     /product="16S ribosomal RNA"
ORIGIN
        1 agagtttgat cctggctcag gacgaacgct ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggcg aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaaaccgg ggctaatacc ggatatgacc atctgtcgca
      181 tggtgggtgg tggaaagatt ttttggcttg ggatgggcta cgcggcctat cagcttgttg
      241 gtggggtgat ggcctaccaa ggcggcgacg ggtagccggc ctgagagggc gaccggccac
      301 actgggactg agacacggcc cagactccta cgggaggcag cagtggggaa tcttgcacaa
      361 tgggcggaag cctgatgcag cgacgccgcg tgagggatga cggccttcgg gttgtaaacc
      421 tctttcagca gggacgaagc gttttgtgacg gtacctgcag aagaagcgcc ggccaactac
      481 gtgccagcag ccgcggtaag acgtagggcg caagcgttgt ccggatttat tgggcgtaaa
      541 gagctcgtag gcggcttgtc gcgtctgactg tgaaaacccg tggctcaact gcgggcttgc
      601 agtcgatacg ggcaggctag agttcggtag gggagactgg aattcctggt gtagcggtga
      661 aatgcgcaga tatcaggagg aacaccggtg gcgaaggcgg gtctctgggc cgatactgac
      721 gctgaggagc gaaagcgtgg ggagcgaaca ggattagata ccctggtagt ccacgctgta
      781 aacgttgggg cgtaggtgtg gggggcctct ccggttctct gtgccgcagc taacgcatta
      841 agcgccccgc ctggggagta cggccgcaag gctaaaactc aaaggaattg acgggggccc
      901 gcacaagcgg cggagcatgc ggattaattc gatgcaacgc gaagaacctt acctgggttt
      961 gacatcgccg gaaatccttc agagatgggg ggtccttcgg ggccggtgac aggtggtgca
     1021 tggctgtcgt cagctcgtgt cgtgagatgt tgggttaagt cccgcaacga gcgcaaccct
     1081 tgttcgatgt tgccagcgcg ttatggcggg gactcatcga agactgccgg ggtcaactcg
     1141 gaggaaggtg gggatgacgt caagtcatca tgccccttat gtccagggct tcacgcatgt
     1201 tacaatggcc ggtacagtgg gctgcgatac cgtgaggtgg agcgaatccc aaaaagccgg
     1261 tctcagttcg gatcggggtc tgcaactcga ccccgtgaag tcggagtcgc tagtaatcgc
     1321 agatcagcaa cgctgcggtg aatacgttcc cgggccttgt acacaccgcc cgtcacgtca
     1381 cgaaagtcgg caacacccga agccggtggc ctaacccttg tgggggggagc cgtcgaaggt
     1441 ggggctggcg attgggacga agtcgtaaca aggtagccgt a
//
```

AY040618 . Salinospora sp. C...[gi:22474392]
Links

```
LOCUS       AY040618                1483 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH536 16S ribosomal RNA gene, partial sequence.
ACCESSION   AY040618
VERSION     AY040618.1  GI:22474392
KEYWORDS    .
SOURCE      Salinospora sp. CNH536
  ORGANISM Salinospora sp. CNH536 Bacteria; Actinobacteria; Actinobacteridae;
Actinomycetales;
            Micromonosporineae; Micromonosporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1483)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.
  TITLE     Widespread and persistent populations of a major new marine
            actinomycete taxon in ocean sediments
  JOURNAL   Appl. Environ. Microbiol. 68 (10), 5005-5011 (2002)
  MEDLINE 22235406 PUBMED 12324350 REFERENCE   2  (bases 1 to 1483)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Institution of
            Oceanography, UCSD, 8602 La Jolla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1483
                     /organism="Salinospora sp. CNH536"
                     /mol_type="genomic DNA"
                     /strain="CNH536"
                     /db_xref="taxon:168696" rRNA <1..>1483
                     /product="16S ribosomal RNA"
ORIGIN
        1 agagtttgat cctggctcag gacgaacgct ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggcg aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaaaccgg ggctaatacc ggatatgact ggctgccgca
      181 tggtggttgg tggaaagatt ttttggcttg ggatgggctc gcggcctatc agcttgttgg
      241 tggggtgatg gcctaccaag gcggcgacgg gtagccggcc tgagagggcg accggccaca
      301 ctgggactga gacacggccc agactcctac gggaggcagc agtggggaat cttgcacaat
      361 gggcggaagc ctgatgcagc gacgccgcgt gagggatgac ggccttcggg ttgtaaacct
      421 ctttcagcag ggacggaagc gttttgtgacg gtacctgcag aagaagcgcc ggccaactac
      481 gtgccagcag ccgcggtaag acgtagggcg caagcgttgt ccggaattta ttgggcgtaa
      541 agagctcgta ggcggcttgt cgcgtcgact gtgaaaaccc gtggctcaac tgcgggcttg
      601 cagtcgatac gggcaggcta gagttcggta ggggagactg gaattcctgg tgtagcggtg
      661 aaatgcgcag atatcaggag gaacaccggt ggcgaaggcg ggtctctggg ccgatactga
      721 cgctgaggag cgaaagcgtg gggagcgaac aggattagat accctggtag tccacgctgt
      781 aaacgttggg cgctaggtgt ggggagcctc tccggttctc tgtgccgcag ctaacgcatt
      841 aagcgccccg cctggggagt acggccgcaa ggctaaaact caaaggaatt gacgggggcc
      901 cgcacaagcg gcggagcatg cggattaatt cgatgcaacg cgaagaacct tacctgggtt
      961 tgacatcgcc ggaaatcctt cagagatggg gggtccttcg gggccggtga caggtggtgc
     1021 atggctgtcg tcagctcgtg tcgtgagatg ttgggttaag tcccgcaacg agcgcaaccc
     1081 ttgttcgatg ttgccagcgc gttatggcgg ggactcatcg aagactgccg gggtcaactc
     1141 ggaggaaggt ggggatgacg tcaagtcatc atgcccctta tgtccagggc ttcacgcatg
     1201 ctacaatggc cggtacaatg ggctgcgata ccgtgaggtg gagcgaatcc caaaaagccg
     1261 gtctcagttc ggatcggggt ctgcaactcg accccgtgaa gtcggagtcg ctagtaatcg
     1321 cagatcagca acgctgcggt gaatacgttc ccgggccttg tacacaccgc ccgtcacgtc
     1381 acgaaagtcg gcaacacccg aagccggtgg ccctaaccct tgtgggggga gccgtcgaag
     1441 gtggggctgg cgaatgggac gaagtcgtaa caaggtagcc gta
//
```

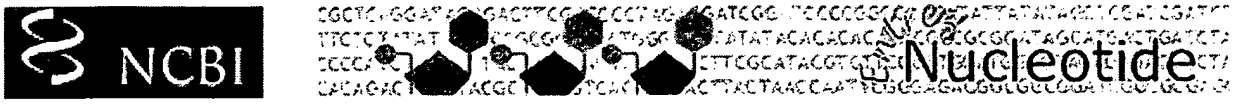
AY040617 . Salinospora sp. C...[gi:22474391]
Links

```
LOCUS       AY040617                1480 bp    DNA     linear   BCT 27-SEP-2002
DEFINITION  Salinospora sp. CNH440 16S ribosomal RNA gene, partial sequence.
ACCESSION   AY040617
VERSION     AY040617.1  GI:22474391
KEYWORDS    .
SOURCE      Salinospora sp. CNH440
  ORGANISM Salinospora sp. CNH440 Bacteria; Actinobacteria; Actinobacteridae;
Actinomycetales;
            Micromonosporineae; Micromonosporaceae; Salinospora.
REFERENCE   1  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.
```

```
  TITLE     Widespread and persistent populations of a major new marine
            actinomycete taxon in ocean sediments
  JOURNAL   Appl. Environ. Microbiol. 68 (10), 5005-5011 (2002)
  MEDLINE 22235406 PUBMED 12324350 REFERENCE   2  (bases 1 to 1480)
  AUTHORS   Mincer,T.J., Jensen,P.R., Kauffman,C.A. and Fenical,W.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Marine Chemistry, Scripps Institution of
            Oceanography, UCSD, 8602 La Jolla Shores Dr., La Jolla, CA
            92093-0204, USA
FEATURES             Location/Qualifiers
     source          1..1480
                     /organism="Salinospora sp. CNH440"
                     /mol_type="genomic DNA"
                     /strain="CNH440"
                     /db_xref="taxon:168695" rRNA <1..>1480
                     /product="16S ribosomal RNA"
ORIGIN
        1 agagtttgat cctggctcag gacgaacgct ggcggcgtgc ttaacacatg caagtcgagc
       61 ggaaaggccc ttcggggtac tcgagcggcg aacgggtgag taacacgtga gtaacctgcc
      121 ccaggctttg ggataacccc gggaaaccgg ggctaatacc ggatatgact ggctgccgca
      181 tggtggttgg tggaaagatt ttttggcttg ggatgggctc gcggcctatc agcttgttgg
      241 tggggtgatg gcctaccaag gcggcgacgg gtagccggcc tgagagggcg accggccaca
      301 ctgggactga gacacggccc agactcctac gggaggcagc agtggggaat cttgcacaat
      361 gggcggaagc ctgatgcagc gacgccgcgt gagggatgac ggccttcggg ttgtaaacct
      421 ctttcagcag ggacgaagcg ttttgtgacgg tacctgcaga agaagcgccg gccaactacg
      481 tgccagcagc cgcggtaaga cgtagggcgc aagcgttgtc cggatttatt gggcgtaaag
      541 agctcgtagg cggcttgtcg cgtcgactgt gaaaacccgt ggctcaactg cgggcttgca
      601 gtcgatacgg gcaggctaga gttcggtagg ggagactgga attcctggtg tagcggtgaa
      661 atgcgcagat atcaggagga acaccggtgg cgaaggcggg tctctgggcc gatactgacg
      721 ctgaggagcg aaagcgtggg gagcgaacag gattagatac cctggtagtc cacgctgtaa
      781 acgttgggcg ctaggtgtgg ggagcctctc cggttctctg tgccgcagct aacgcattaa
      841 gcgccccgcc tggggagtac ggccgcaagg ctaaaactca aaggaattga cgggggcccg
      901 cacaagcggc ggagcatgcg gattaattcg atgcaacgcg aagaacctta cctgggtttg
      961 acatcgccgg aaatccttca gagatggggg gtccttcggg gccggtgaca ggtggtgcat
     1021 ggctgtcgtc agctcgtgtc gtgagatgtt gggttaagtc ccgcaacgag cgcaaccctt
     1081 gttcgatgtt gccagcgcgt tatggcgggg actcatcgaa gactgccggg gtcaactcgg
     1141 aggaaggtgg ggatgacgtc aagtcatcat gccccttatg tccagggctt cacgcatgct
     1201 acaatggccg gtacaatggg ctgcgatacc gtgaggtgga gcgaatccca aaaagccggt
     1261 ctcagttcgg atcggggtct gcaactcgac cccgtgaagt cggagtcgct agtaatcgca
     1321 gatcagcaac gctgcggtga atacgttccc gggccttgta cacaccgccc gtcacgtcac
     1381 gaaagtcggc aacacccgaa gccggtggcc taacccttgt ggggggagcc gtcgaaggtg
     1441 gggctggcga ttgggacgaa gtcgtaacaa ggtagccgta
//
```

**EXHIBIT G**

Genbank search results to identify CNH964 Genbank entry






About Entrez

Search for Genes
LocusLink provides curated information for human, fruit fly, mouse, rat, and zebrafish

Entrez Nucleotide
Help | FAQ

Batch Entrez: Upload a file of GI or accession numbers to retrieve sequences

Check sequence revision history

How to create WWW links to Entrez

LinkOut

Cubby

Related resources
BLAST

Reference sequence project

Submit to GenBank

Display Summary Show: 20 Send to Text

Items 1-2 of 2 One page.

1: AY371900 Links
Salinospora sp. CNH964 16S-23S ribosomal RNA intergenic spacer
gi|34334163|gb|AY371900.1|[34334163]

2: AY040623 Links
Salinospora sp. CNH964 16S ribosomal RNA gene, partial sequence
gi|22474397|gb|AY040623.1|[22474397]



Oct 14 2003 07:20:40